

# Convergence of Proximal Gradient Algorithm in the Presence of Adjoint Mismatch<sup>‡</sup>

Emilie Chouzenoux<sup>1</sup>, Jean-Christophe Pesquet<sup>1</sup>, Cyril Riddell<sup>2</sup>, Marion Savanier<sup>1,2,§</sup>, and Yves Troussel<sup>2</sup>

<sup>1</sup>: University of Paris-Saclay, CentraleSupélec, CVN, Inria, Gif-sur-Yvette, France

<sup>2</sup>: GE Healthcare, Buc, France

**Abstract.** We consider the proximal gradient algorithm for solving penalized least-squares minimization problems arising in data science. This first-order algorithm is attractive due to its flexibility and minimal memory requirements allowing to tackle large-scale minimization problems involving non-smooth penalties. However, for problems such as X-ray computed tomography, the applicability of the algorithm is dominated by the cost of applying the forward linear operator and its adjoint at each iteration. In practice, the adjoint operator is thus often replaced by an alternative operator with the aim to reduce the overall computation burden and potentially improve conditioning issues. In this paper, we propose to analyze the effect of such an adjoint mismatch on the convergence of the proximal gradient algorithm in an infinite-dimensional setting, thus generalizing the existing results on PGA. We derive conditions on the step-size and on the gradient of the smooth part of the objective function under which convergence of the algorithm to a fixed point is guaranteed. We also derive bounds on the error between this point and the solution to the original minimization problem. We illustrate our theoretical findings with two image reconstruction tasks in computed tomography.

**Keywords:** Adjoint mismatch, convex optimization, computed tomography, convergence analysis, fixed point methods, image reconstruction, forward-backward algorithm

## 1. Introduction

Linear inverse problems arise when modeling phenomena from a broad range of real-life applications in image and signal processing. They aim to recover an estimate of an unknown signal  $\bar{x}$ , assumed to belong to a real Hilbert space  $\mathcal{H}$ , from the following observation model

$$y = H\bar{x} + b, \tag{1}$$

where  $y$  is the known measurement assumed to belong to a real Hilbert space  $\mathcal{G}$ ,  $H$  is a bounded linear operator from  $\mathcal{H}$  to  $\mathcal{G}$ , and  $b \in \mathcal{G}$  is a noise term. Such a problem is

<sup>‡</sup> This work was supported by the European Research Council Starting Grant MAJORIS ERC-2019-STG-850925, the ANRT CIFRE Convention 2018/1587, the ANR AI Chair BRIGEABLE, and the Institut Universitaire de France.

<sup>§</sup> Corresponding author. Contact: marion.savanier@centralesupelec.fr

generally ill-posed as  $H$  is poorly conditioned thus making the solution oversensitive to perturbations in the measurements. In this context, the Moore-Penrose pseudo inversion of  $H$  does not produce a meaningful solution. To avoid these shortcomings, variational methods propose to provide an estimate  $x \in \mathcal{H}$  of  $\bar{x}$ , by minimizing a composite objective function summing a data fidelity term and a penalization term. The data fidelity term measures the discrepancy between the observation  $y$  and its estimate  $Hx$  for a given statistical model on the noise. In the following, we focus on the additive white Gaussian noise, so that the natural data fidelity is quadratic and its minimum is given by the Moore-Penrose generalized inverse. The penalization term shifts the solution away from this rough estimate by adding a priori information about the sought solution.

In most modern scientific areas, the availability of vast amounts of data has led to large-scale inverse problems for which a minimizer must be processed quickly. To this end, first order optimization algorithms are effective tools because they can be conveniently implemented and do not require onerous computations. Among them is the proximal gradient algorithm (PGA) [17], itself an instance of the forward-backward algorithm [20]. It consists of an explicit gradient descent step which exploits the differentiability and convexity of the data fidelity term and an implicit proximity step, which only assumes the convexity of the penalization term.

A quadratic fidelity term involves the positive self adjoint operator  $H^*H$ , where  $H^*$  denotes the adjoint of  $H$ . Actually, only a product  $H^*Hx$ , with  $x \in \mathcal{H}$ , is required when evaluating its gradient at each iteration. Designing efficient implementations of this product is thus key to ensure speed and applicability of PGA. Hence the adjoint  $H^*$  is often replaced by an alternative operator, with the aim to increase the convergence rate, thanks to better conditioning, or to make efficient use of hardware accelerators and therefore alleviate the total computation time.

Although this strategy results in an adjoint mismatch that breaks the operator symmetry [29, 49], it is frequently used in tomographic transmission imaging [29], as practiced in industrial non-destructive testing and diagnostic medical imaging [8, 33] and it has also been advocated in SPECT (Single Photon Emission Computed Tomography) imaging [44, 50]. In tomography,  $\bar{x}$  represents a scanned object with a spatially varying property that the forward projection operator  $H$  maps to a set of projections corresponding to measurements at different angles. The adjoint operator  $H^*$  is then the backprojector. Both projection and backprojection operations act in the continuous domain and depend on the physical properties of the object. When discretizing them on a Cartesian grid, different strategies can be used. In SPECT, the modeling of the attenuation may be bypassed in the backprojector. In X-ray tomography, due to the changes in sampling rates induced by the rotation of the system and the divergent geometry of the X-ray beam, oversampling in the image domain is necessary to derive an accurate projector. However it implies redundancies that are usually unwanted and thus removed in the backprojector. Popular GPU implementations of these projection and backprojection operators are also often unmatched as it is the case in the ASTRA software package [47].

State-of-the-art matched projection/backprojection pairs rely on advanced interpolation schemes that are difficult to parallelize on GPU [11, 37, 38]. Using an unmatched pair to reduce the computational costs however endangers the convergence of the reconstruction algorithm. It is an open problem to know whether errors will accumulate over iterations [1, 49] and lead to suboptimal reconstruction compared to having a matched pair.

The literature remains scarce on the influence of such mathematical approximation on the convergence of PGA. When the objective function is a least-squares term without any regularization, PGA is reduced to a simple gradient algorithm. In this context, adjoint mismatch has been investigated in the early work of [50] and then in [24, 27, 39]. In [27], the authors studied adjoint mismatch and additional perturbations of their algebraic scheme. Later, in [24], they discussed a modification of the algorithm that can guarantee convergence to a slightly different problem in finite dimension. Their analysis can be extended in a direct manner to the case when the regularization term is a quadratic function. The work in [39] also dealt with a least-squares objective function but using projections onto convex sets in the randomized Kaczmarz method. Studying the impact of adjoint mismatch also finds application in an active line of research, namely deep learning. Very recently, Bubba *et al.* have proposed a CNN-based reconstruction algorithm  $\Phi$ DONet in [7] where the backprojection operator is replaced by a partially learned approximation to it thanks to an unrolled ISTA (Iterative Soft-Thresholding Algorithm, i.e. a special case of PGA) structure, with the aim to improve the backprojection in the context of limited angle tomography. In that case, the considered prior was an  $\ell_1$  penalization applied to the wavelets coefficients of the object. Under some conditions on  $H$ ,  $\Phi$ DONet can be seen as a perturbed version of ISTA. By adopting a probabilistic approach, the authors establish the convergence in mean of the output of their optimally trained network with respect to the ground truth in finite dimension. Note that their approach, though targeted to a specific application, could be extended to any convolutional forward operator  $H$  which is a pseudodifferential operator or a Fourier integral operator.

Up to our knowledge, PGA stability in the presence of adjoint mismatch has not been considered so far when general nonlinear operators induced by the presence of potentially non differentiable convex priors are involved. In this paper, we propose to extend the theoretical ideas in [24] on PGA in the presence of adjoint mismatch to solve a penalized least-squares problem in an arbitrary Hilbert space. For this kind of problems, the resulting algorithm can be seen as a generalization of PGA.

The major contributions of this paper are:

- a characterization of the fixed points of PGA in the presence of an adjoint mismatch;
- conditions of convergence with new bounds on the gradient step-size and on the regularization parameters;
- a characterization of the distance from the generated fixed point of the algorithm to a “true” minimizer of the original objective function;
- a validation of these results on image reconstruction scenarios.

The paper is organized as follows: Section 2 introduces the notation used in this work and recalls results about the proximal gradient. Section 3 gives necessary conditions to preserve the convergence of PGA with an adjoint mismatch and gives a bound on the discrepancy induced by the mismatch on the resulting fixed point and the minimizer of the original objective function. Next, examples of linear inverse problems with sparsity constraints arising from computed tomography are discussed in Section 4. Finally, some conclusions are drawn in Section 5.

## 2. Preliminaries

### 2.1. Notation and mathematical background

We first provide the necessary background on convex analysis, proximity operators and monotone operators. In this paper, the underlying signal space is the real Hilbert space  $\mathcal{H}$  endowed with scalar product  $\langle \cdot | \cdot \rangle$ , norm  $\| \cdot \|$  and identity operator  $\text{Id}$ .  $2^{\mathcal{H}}$  denotes the power set of  $\mathcal{H}$ .  $\mathcal{B}(\mathcal{H}, \mathcal{G})$  denotes the Banach space of bounded linear operators from  $\mathcal{H}$  to  $\mathcal{G}$ , equipped with norm  $\| \cdot \|_{\mathcal{S}}$  so as

$$(\forall T \in \mathcal{B}(\mathcal{H}, \mathcal{G})) \quad \|T\|_{\mathcal{S}} = \sup_{\substack{x \in \mathcal{H} \\ \|x\| \leq 1}} \|Tx\|. \quad (2)$$

The adjoint of  $T$  is denoted by  $T^* \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ . Moreover,  $\text{Ker } T$  designates the nullspace of operator  $T$  and  $\text{ran } T$  its range. If  $\text{ran } T$  is closed, its pseudo-inverse is  $T^\# \in \mathcal{B}(\mathcal{G}, \mathcal{H})$ .

The class of functions which are proper, convex, lower-semicontinuous on  $\mathcal{H}$  and take values in  $\mathbb{R} \cup \{+\infty\}$  is denoted by  $\Gamma_0(\mathcal{H})$ . For every  $g \in \Gamma_0(\mathcal{H})$ ,  $\text{dom } g$  is the domain of  $g$  and  $\partial g$  is the subdifferential of  $g$  [2]. If  $x \in \mathcal{H}$ , the proximity operator of  $g$  at  $x$  is defined as [40]

$$\text{prox}_g(x) = \underset{z \in \mathcal{H}}{\text{argmin}} \left( g(z) + \frac{1}{2} \|x - z\|^2 \right). \quad (3)$$

We say that  $f \in \Gamma_0(\mathcal{H})$  is coercive if

$$\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty \quad (4)$$

and supercoercive if

$$\lim_{\|x\| \rightarrow +\infty} \frac{f(x)}{\|x\|} = +\infty. \quad (5)$$

We shall also require the following properties of cocoercivity and monotonicity.

An operator  $A : \mathcal{H} \mapsto 2^{\mathcal{H}}$  is said to be monotone if

$$(\forall (x, y) \in \mathcal{H}^2)(\forall u \in Ax)(\forall v \in Ay) \quad \langle u - v | x - y \rangle \geq 0. \quad (6)$$

It is said to be maximal monotone if, in addition, its graph  $\{(x, y) \in \mathcal{H}^2 \mid y \in Ax\}$ , is not properly contained in the graph of any other monotone operator.  $A : \mathcal{H} \mapsto 2^{\mathcal{H}}$  is said to be strictly monotone if

$$(\forall (x, y) \in \mathcal{H}^2)(\forall u \in Ax)(\forall v \in Ay) \quad x \neq y \Rightarrow \langle u - v | x - y \rangle > 0. \quad (7)$$

$A : \mathcal{H} \mapsto 2^{\mathcal{H}}$  is said to be strongly monotone if there exist  $\eta \in ]0, +\infty[$  such that  $A - \eta \text{Id}$  is monotone. A key property to ensure the convergence of PGA is the cocoercivity of the involved gradient operator or of its approximation. Let us recall that operator  $A : \mathcal{H} \rightarrow \mathcal{H}$  is  $\eta$ -cocoercive with  $\eta \in [0, +\infty[$  if

$$(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \quad \eta \|Ax - Ay\|^2 \leq \langle x - y | Ax - Ay \rangle. \quad (8)$$

Moreover,  $A$  is said to be nonexpansive if it is Lipschitz continuous with constant 1, i.e.,

$$(\forall (x, y) \in \mathcal{H}^2) \quad \|Ax - Ay\| \leq \|x - y\|. \quad (9)$$

$A$  is  $\alpha$ -averaged with  $\alpha \in ]0, 1]$  if there exists a nonexpansive operator  $Q$  such that  $A = (1 - \alpha)\text{Id} + \alpha Q$ . If  $A : \mathcal{H} \rightarrow \mathcal{H}$ ,  $\text{Fix } A$  denotes the set of its fixed points.

## 2.2. Proximal gradient algorithm for the penalized least-squares criterion

PGA is used to find an estimate of  $\bar{x}$ , defined in (1) by solving the following penalized least squares criterion:

$$\underset{x \in \mathcal{H}}{\text{minimize}} \quad \frac{1}{2} \|y - Hx\|^2 + g(x) + \frac{\kappa}{2} \|x\|^2, \quad (10)$$

where  $g \in \Gamma_0(\mathcal{H})$  is a suitable possibly non-smooth regularization function and  $\kappa \in [0, +\infty[$ . Note that (10) covers a wide range of penalization strategies. For instance, when  $\kappa > 0$ , the objective function in (10) is strongly convex and an elastic net-like penalization is thus introduced [54]. Another popular example is when  $\kappa = 0$  and  $g$  is a (possibly weighted)  $\ell_1$  norm defined on a suitable Hilbert space (e.g.  $\mathcal{H} = \mathbb{R}^N$ ) serving to promote the sparsity of the resulting solution in some transform domain [16].

For optimization problem (10), PGA reads, for every  $n \in \mathbb{N}$ ,

$$x_{n+1} = x_n + \theta_n (\text{prox}_{\gamma g}((1 - \gamma\kappa)x_n - \gamma H^*(Hx_n - y)) - x_n), \quad (11)$$

where  $x_0 \in \mathcal{H}$  is the initial estimate,  $(\theta_n)_{n \in \mathbb{N}}$  are nonnegative relaxation parameters, and  $\gamma \in ]0, +\infty[$  is the algorithm step-size. When a sparsity inducing penalty is embedded as mentionned above, this algorithm reduces to the well-known ISTA that was developed for the purpose of wavelet-based signal restoration [3, 10, 22, 28] and later extended to other regularization functions [12].

If  $\theta_n \in [\epsilon, 1]$  with  $\epsilon \in ]0, 1[$  and  $\gamma \in ]0, 2/(\|H\|_S^2 + \kappa)[$ , the sequence  $(x_n)_{n \in \mathbb{N}}$  generated by algorithm (11) converges weakly to a solution to Problem (10) when such a solution exists [16, 20, 22]. Strong convergence is even achieved in some contexts [4, 16, 22]. Recent results on overrelaxed versions of (11) can be found in [21] for special cases of gradient operators. The flexibility introduced by an iteration-dependent parameters step-size can be used to improve the algorithm convergence pattern. Without loss of generality, the step-size will be hereafter assumed to be constant knowing that extending our analysis to varying step-sizes is straightforward.

## 3. Convergence analysis with an adjoint mismatch

### 3.1. Mismatched algorithm

As mentioned earlier, in the context of an adjoint mismatch, operator  $H^*$  is purposefully replaced by surrogate operators  $(K_n)_{n \in \mathbb{N}}$ , iteration (11) thus becoming: For every  $n \in \mathbb{N}$ ,

$$x_{n+1} = x_n + \theta_n (\text{prox}_{\gamma g}((1 - \gamma\kappa)x_n - \gamma K_n(Hx_n - y)) - x_n). \quad (12)$$

Hereafter, we list assumptions used throughout this paper, to analyze scheme (12).

#### Assumption 3.1

- (i)  $g \in \Gamma_0(\mathcal{H})$
- (ii) For every  $n \in \mathbb{N}$ ,  $K_n \in \mathcal{B}(\mathcal{G}, \mathcal{H})$
- (iii) There exist  $\bar{K} \in \mathcal{B}(\mathcal{G}, \mathcal{H})$  and  $\{\omega_n\}_{n \in \mathbb{N}} \subset ]0, +\infty[$  with  $\sum_{n \in \mathbb{N}} \omega_n < +\infty$  such that

$$\bar{K}H \neq 0 \quad (13)$$

$$(\forall n \in \mathbb{N}) \quad \|K_n - \bar{K}\|_S \leq \omega_n. \quad (14)$$

The last assumption covers two scenarios of particular interest:

- When  $\bar{K} = H^*$ , we get a sequence of approximations  $(K_n)_{n \in \mathbb{N}}$  providing asymptotically a perfect estimation of the adjoint of  $H$ .
- When, for every  $n \in \mathbb{N}$ ,  $\omega_n = 0$ , a constant error  $\bar{K} - H^*$  is introduced on the adjoint.

In the context of convergence analysis of fixed point iterations [19] of the modified PGA algorithm (12), the following notation is central.

**Notation 3.2** Let  $\gamma \in ]0, +\infty[$ . We define

$$L = \bar{K}H + \kappa \text{Id} \quad (15)$$

$$\begin{aligned} T_\gamma: \mathcal{H} &\rightarrow \mathcal{H} \\ x &\mapsto \text{prox}_{\gamma g}(x - \gamma Lx + \gamma \bar{K}y) \end{aligned} \quad (16)$$

$$\begin{aligned} \lambda_{\min} &= \inf_{\substack{x \in \mathcal{H} \\ \|x\|=1}} \langle x | Lx \rangle \\ \lambda_{\min}^+ &= \inf_{\substack{x \in (\text{Ker } L)^\perp \\ \|x\|=1}} \langle x | Lx \rangle \\ \lambda_{\max} &= \sup_{\substack{x \in \mathcal{H} \\ \|x\|=1}} \langle x | Lx \rangle \\ \beta &= \frac{1}{2} \|L - L^*\|_{\text{S}}. \end{aligned} \quad (17)$$

Note that  $\lambda_{\min}$  (resp.  $\lambda_{\max}$ ) is the minimum (resp. maximum) spectral value of  $(L + L^*)/2$  and that  $\lambda_{\min}^+ \geq \lambda_{\min}$ . We will show that the convergence of algorithm (12) is guaranteed under cocoercivity conditions on the operator  $L$ .

### 3.2. Properties of the modified gradient descent operator

When  $\bar{K} \neq H^*$ , the gradient of the smooth part of our objective function is replaced by the operator  $\kappa \text{Id} + \bar{K}(H \cdot - y)$ , which is not guaranteed to be a cocoercive operator. We will thus propose conditions preserving this property. First, we prove certain properties induced by the cocoercivity of operator  $L$ , which will be used throughout the article.

**Lemma 3.3** Let  $\eta \in ]0, +\infty[$ . If  $L$  is  $\eta$ -cocoercive, then the following hold:

- (i)  $\lambda_{\min} \geq 0$
- (ii)  $\text{Ker}(L + L^*) = \text{Ker } L = \text{Ker } L^*$
- (iii)  $L + L^* \neq 0$ .

*Proof.*  $L$  is  $\eta$ -cocoercive if and only if, for every  $x \in \mathcal{H}$ ,

$$\langle x | Lx \rangle \geq \eta \|Lx\|^2. \quad (18)$$

(i): The fact that  $\lambda_{\min} \geq 0$  directly follows from (18).

(ii): If  $x \in \text{Ker } L$ , then

$$\begin{aligned} \langle x | Lx \rangle &= 0 \\ \Leftrightarrow \langle x | (L + L^*)x \rangle &= 0 \end{aligned} \quad (19)$$

According to (i),  $L + L^*$  is self-adjoint positive. It thus admits a self adjoint square root  $(L + L^*)^{1/2}$  and (19) is equivalent to

$$\|(L + L^*)^{1/2}x\|^2 = 0 \quad \Leftrightarrow \quad (L + L^*)^{1/2}x = 0, \quad (20)$$

which yields  $(L + L^*)x = 0$ . We have thus proved that  $\text{Ker } L \subset \text{Ker } (L + L^*)$ . By reexpressing (18),

$$(\forall x \in \mathcal{H}) \quad \frac{1}{2} \langle x \mid (L + L^*)x \rangle \geq \eta \|Lx\|^2. \quad (21)$$

Consequently, if  $x \in \text{Ker } (L + L^*)$ , then  $x \in \text{Ker } L$ . In summary,  $\text{Ker } (L + L^*) = \text{Ker } L$ . By symmetry,  $\text{Ker } (L + L^*) = \text{Ker } L^*$ .

(iii):  $L + L^* = 0$  if and only if  $\text{Ker } (L + L^*) = \mathcal{H}$  which, according to (ii), would imply that  $\text{Ker } L = \mathcal{H}$ , that is  $L = 0$ . This contradicts our assumption in (13).  $\square$

Whenever cocoercivity is present, we will show that the behavior of iterative scheme (12) remains stable. Conditions for cocoercivity are summarized below.

**Proposition 3.4**

(i) Assume that  $\lambda_{\min} \geq 0$ .

If  $\lambda_{\min}^+ \in ]0, +\infty[$  and  $\text{Ker } (L + L^*) = \text{Ker } L$ , then  $L$  is  $\underline{\eta}$ -cocoercive with

$$\underline{\eta} = 1 / \left( \sqrt{\lambda_{\max}} + \frac{\beta}{\sqrt{\lambda_{\min}^+}} \right)^2. \quad (22)$$

If  $\beta = 0$ , then  $L$  is  $(1/\lambda_{\max})$ -cocoercive.

(ii) Suppose that  $\text{ran } (L + L^*)$  is closed.  $L$  is  $\eta$ -cocoercive with  $\eta \in ]0, +\infty[$  if and only if  $\lambda_{\min} \geq 0$ ,  $\text{Ker } (L + L^*) = \text{Ker } L$ , and

$$\eta \leq \bar{\eta} = \frac{2}{\|(\text{Id} + (L - L^*)(L + L^*)^\#)(L + L^*)^{1/2}\|_S^2}. \quad (23)$$

*Proof:* (i): Let  $A$  and  $B$  be the self-adjoint and skewed parts of  $L$ , respectively given by

$$A = \frac{L + L^*}{2} \quad (24)$$

$$B = \frac{L - L^*}{2}. \quad (25)$$

Assume first that  $V = \text{Ker } A = \text{Ker } L = \text{Ker } L^*$ . Let  $x \in \mathcal{H}$  and let  $x_{V^\perp}$  denote its projection onto the orthogonal complement of  $V$ . We have

$$\begin{aligned} \|Lx\|^2 &= \|Lx_{V^\perp}\|^2 \\ &\leq (\|Ax_{V^\perp}\| + \|Bx_{V^\perp}\|)^2. \end{aligned} \quad (26)$$

Since  $\lambda_{\min} \geq 0$ ,  $A$  is a positive operator and we have then

$$\|Ax_{V^\perp}\|^2 \leq \|A\|_S \langle x_{V^\perp} \mid Ax_{V^\perp} \rangle = \lambda_{\max} \langle x_{V^\perp} \mid Ax_{V^\perp} \rangle. \quad (27)$$

In turn,

$$\|Bx_{V^\perp}\| \leq \beta \|x_{V^\perp}\|. \quad (28)$$

and

$$\langle x_{V^\perp} \mid Ax_{V^\perp} \rangle \geq \lambda_{\min}^+ \|x_{V^\perp}\|^2. \quad (29)$$

Then, if  $\lambda_{\min}^+ > 0$ ,

$$\|Bx_{V^\perp}\|^2 \leq \frac{\beta^2}{\lambda_{\min}^+} \langle x_{V^\perp} \mid Ax_{V^\perp} \rangle \quad (30)$$

Altogether (26), (27) and (30) yield

$$\|Lx\|^2 \leq \frac{1}{\underline{\eta}} \langle x_{V^\perp} | Ax_{V^\perp} \rangle = \frac{1}{\underline{\eta}} \langle x | Ax \rangle = \frac{1}{\underline{\eta}} \langle x | Lx \rangle, \quad (31)$$

where  $\underline{\eta}$  is given by (22). This shows that  $L$  is  $\underline{\eta}$ -cocoercive.

If  $\beta = 0$ , then  $L = A$  and the result follows from the inequality

$$\|Ax\|^2 \leq \lambda_{\max} \langle x | Ax \rangle. \quad (32)$$

(ii): According to Lemma 3.3, if  $L$  is cocoercive then  $\lambda_{\min} \geq 0$  and  $\text{Ker}(L + L^*) = \text{Ker } L$ . To establish the result, we will thus assume that these two conditions are satisfied and prove that  $L$  is cocoercive if and only if (23) holds.

Let us use the same notation as in the proof of (i). Since  $V = \text{Ker } A = \text{Ker } L$ ,  $L$  is  $\eta$ -cocoercive with  $\eta \in ]0, +\infty[$  if and only if

$$(\forall x \in V^\perp) \quad \eta \|Lx\|^2 \leq \langle x | Ax \rangle. \quad (33)$$

Let  $x \in V^\perp$  and let  $y = Ax$ . Since  $\text{ran } A$  is closed, this is equivalent to  $x = A^\# y$ . We have then

$$\begin{aligned} \|Lx\|^2 &= \|Ax + Bx\|^2 \\ &= \|Ax + BA^\# y\|^2 \\ &= \|(\text{Id} + BA^\#)Ax\|^2 \\ &\leq \|(\text{Id} + BA^\#)A^{1/2}\|_S^2 \|A^{1/2}x\|^2 \\ &= \|(\text{Id} + BA^\#)A^{1/2}\|_S^2 \langle x | Ax \rangle. \end{aligned} \quad (34)$$

Note that  $\|(\text{Id} + BA^\#)A^{1/2}\|_S \neq 0$  (since  $L$  is nonzero). We have thus shown that  $L$  is cocoercive with constant  $1/\|(\text{Id} + BA^\#)A^{1/2}\|_S^2 = \bar{\eta}$ , hence for any constant  $\eta > 0$  satisfying (23).

In addition, the maximum cocoercity constant  $\eta_{\max}$  is such that

$$\frac{1}{\eta_{\max}} = \sup_{v \in V^\perp \setminus \{0\}} \frac{\|Lx\|^2}{\langle x | Ax \rangle} = \sup_{v \in V^\perp \setminus \{0\}} \frac{\|(\text{Id} + BA^\#)Ax\|^2}{\|A^{1/2}x\|^2}. \quad (35)$$

On the other hand,

$$\|(\text{Id} + BA^\#)A^{1/2}\|_S^2 = \sup_{z \in \mathcal{H} \setminus \{0\}} \frac{\|(\text{Id} + BA^\#)A^{1/2}z\|^2}{\|z\|^2}. \quad (36)$$

Every  $z \in \mathcal{H}$  can be decomposed as  $z_V + z_{V^\perp}$ , where  $(z_V, z_{V^\perp}) \in V \times V^\perp$ . Since  $A$  is self-adjoint positive,  $V = \text{Ker } A^{1/2} = \text{Ker } A$ . We can thus reexpress (36) as

$$\begin{aligned} \|(\text{Id} + BA^\#)A^{1/2}\|_S &= \sup_{z \in \mathcal{H} \setminus \{0\}} \frac{\|(\text{Id} + BA^\#)A^{1/2}z_{V^\perp}\|^2}{\|z_V\|^2 + \|z_{V^\perp}\|^2} \\ &= \sup_{z_{V^\perp} \in V^\perp \setminus \{0\}} \frac{\|(\text{Id} + BA^\#)A^{1/2}z_{V^\perp}\|^2}{\|z_{V^\perp}\|^2}. \end{aligned} \quad (37)$$

We know however that  $V^\perp = (\text{Ker } A^{1/2})^\perp = \overline{\text{ran}((A^{1/2})^*)} = \overline{\text{ran}(A^{1/2})}$  []. The expressions in (35) and (37) are thus equal.  $\square$

**Remark 3.5** When  $\beta \neq 0$ , (22) suggests that  $\eta$  is higher when the nonzero spectral values of  $(L + L^*)/2$  are clustered together.



The following special cases are worth being mentioned.

**Corollary 3.6**

(i) If  $\lambda_{\min} > 0$ , then  $L$  is cocoercive with constant

$$\begin{aligned}\bar{\eta} &= \frac{2}{\|(\text{Id} + (L - L^*)(L + L^*)^{-1})(L + L^*)^{1/2}\|_S^2} \\ &\geq 1/\left(\sqrt{\lambda_{\max}} + \frac{\beta}{\sqrt{\lambda_{\min}}}\right)^2.\end{aligned}\quad (38)$$

(ii) Assume that  $\mathcal{H}$  is finite dimensional and  $\lambda_{\min} \geq 0$ . If the dimensions of  $\text{Ker } L$  and  $\text{Ker } (L + L^*)$  are equal, then  $L$  is cocoercive with constant  $\bar{\eta} \geq \underline{\eta}$  where  $\underline{\eta}$  and  $\bar{\eta}$  are given by (22) and (23), respectively.

*Proof.*

- (i) If  $\lambda_{\min} > 0$ , then  $L + L^*$  is strongly positive. It is thus invertible,  $\text{ran } (L + L^*) = \mathcal{H}$  is closed, and  $\text{Ker } (L + L^*) = \{0\}$ . In the proof of Proposition 3.4(ii), we have seen that  $\text{Ker } L \subset \text{Ker } (L + L^*)$ . Therefore,  $\text{Ker } (L + L^*)$  and  $\text{Ker } L$  reduce to the null space and, according to Proposition 3.4(ii),  $L$  is  $\bar{\eta}$ -cocoercive. In addition, in this case,  $\lambda_{\min}^+ = \lambda_{\min}$ , it then follows from Proposition 3.4(i) and the fact that  $\bar{\eta}$  is the maximum cocoercivity constant of  $L$  that the lower bound in (38) holds.
- (ii) We have seen that  $\lambda_{\min} \geq 0$  implies that  $\text{Ker } L \subset \text{Ker } (L + L^*)$ . Therefore,  $\text{Ker } L$  is equal to  $\text{Ker } (L + L^*)$  if and only if the dimensions of  $\text{Ker } L$  and  $\text{Ker } (L + L^*)$  are equal. In addition,  $\text{ran } (L + L^*)$  is closed and  $\lambda_{\min}^+$  is necessarily positive when  $\mathcal{H}$  is finite dimensional. The result then follows from Proposition 3.4.

□

**Remark 3.7** Let  $\tilde{\lambda}_{\min}$  be the minimum spectral value of  $(\overline{K}H + H^*\overline{K}^*)/2$ . We have  $\lambda_{\min} = \tilde{\lambda}_{\min} + \kappa$ . A practical choice for  $\kappa$  to ensure that  $\lambda_{\min}$  is positive is thus to set  $\kappa > -\tilde{\lambda}_{\min}$ .

**Remark 3.8** A characterization of cocoercive linear operators in finite dimension was provided in [53] through psd-plus matrices. The characterization provided in this paper may appear more relevant to the context of inverse problems and is valid in an infinite dimensional setting.

To ensure the convergence of (12), it might appear more natural to rely upon conditions which are based on the nonexpansiveness of  $\text{Id} - \gamma L$ . We next show that such conditions are directly related to the cocoercivity of  $L$ .

**Proposition 3.9** If  $L$  is  $\eta$ -cocoercive with  $\eta \in ]0, +\infty[$  and  $\gamma \leq 2\eta$ , then

$$\|\text{Id} - \gamma L\|_S^2 \leq 1 + \gamma\left(\frac{\gamma}{\eta} - 2\right)\lambda_{\min} \leq 1. \quad (39)$$

Conversely, if  $\|\text{Id} - \gamma L\|_S \leq 1$  for some  $\gamma \in ]0, +\infty[$ , then  $L$  is  $\eta$ -cocoercive for every  $\eta \in ]0, \gamma/2]$ .

*Proof:* For every  $x \in \mathcal{H}$ ,

$$\|(\text{Id} - \gamma L)x\|^2 = \|x\|^2 - 2\gamma\langle x | Lx \rangle + \gamma^2\|Lx\|^2. \quad (40)$$

Because of (18)

$$\|(\text{Id} - \gamma L)x\|^2 \leq \|x\|^2 - 2\gamma\langle x | Lx \rangle + \frac{\gamma^2}{\eta}\langle x | Lx \rangle. \quad (41)$$

Therefore, since  $\gamma \leq 2\eta$ ,

$$\|\text{Id} - \gamma L\|_S^2 \leq 1 + \gamma \left( \frac{\gamma}{\eta} - 2 \right) \lambda_{\min}. \quad (42)$$

According to Proposition 3.4(i),  $\lambda_{\min} \geq 0$  and the obtained upper bound is thus less than or equal to 1.

Conversely, if  $\text{Id} - \gamma L$  is nonexpansive, then

$$(\forall x \in \mathcal{H}) \quad \|x - \gamma Lx\|^2 \leq \|x\|^2 \quad (43)$$

The cocoercivity of  $L$  thus straightforwardly follows from (40).  $\square$

### 3.3. Fixed points

As we will see, the fixed point set of operator  $T_\gamma$  plays a prominent role in the convergence analysis of the mismatched PGA. The next proposition characterizes the existence and uniqueness of such a fixed point. Note that a fixed point will generally no longer coincide with the global solution to (10).

#### Proposition 3.10

(i) Let  $\gamma \in ]0, +\infty[$  and let  $\tilde{x} \in \mathcal{H}$ . We have  $\tilde{x} \in \text{Fix } T_\gamma$  if and only if  $\tilde{x}$  belongs to

$$\mathcal{F} = \{x \in \mathcal{H} \mid 0 \in Lx - \overline{K}y + \partial g(x)\}. \quad (44)$$

In addition,  $\mathcal{F}$  is non empty if  $L + \partial g$  is surjective.

(ii) If  $\lambda_{\min} \geq 0$ , then  $\mathcal{F}$  is a closed and convex set.

(iii)  $\mathcal{F}$  has at most one element if one of the following conditions holds:

- (a)  $L + \partial g$  is strictly monotone;
- (b)  $L + L^*$  is positive definite;
- (c)  $\lambda_{\min} \geq 0$  and  $g$  is strictly convex.

In addition,  $\mathcal{F}$  is a singleton if  $\lambda_{\min} \geq 0$  and one of the following conditions holds:

- (d)  $L + \partial g$  is strongly monotone;
- (e)  $\lambda_{\min} \neq 0$ ;
- (f)  $g$  is strongly convex.

(iv) Assume that  $L$  is cocoercive.  $\mathcal{F}$  is nonempty if one of the following conditions holds:

- (a)  $\text{dom } \partial g = \mathcal{H}$  and  $(L + L^*)/2 + \partial g$  is surjective;
- (b)  $\text{dom } \partial g = \mathcal{H}$  and

$$x \mapsto \frac{1}{2} \langle x \mid Lx \rangle + g(x) \quad (45)$$

is coercive;

- (c)  $g$  is supercoercive;
- (d)  $\text{dom } g$  is bounded.

*Proof:* (i): We have

$$\begin{aligned} \tilde{x} \in \text{Fix } T_\gamma &\Leftrightarrow \tilde{x} = \text{prox}_{\gamma g}((1 - \gamma\kappa)\tilde{x} - \gamma\overline{K}(H\tilde{x} - y)) \\ &\Leftrightarrow (1 - \gamma\kappa)\tilde{x} - \gamma\overline{K}(H\tilde{x} - y) \in (\text{Id} + \gamma\partial g)(\tilde{x}) \\ &\Leftrightarrow \tilde{x} \in \mathcal{F}. \end{aligned} \quad (46)$$

Under the considered surjectivity condition, there straightforwardly exists  $\tilde{x} \in \mathcal{H}$  for which (44) holds.

(ii): If  $\lambda_{\min} \geq 0$ , then  $L$  is monotone. Since it is continuous, it is maximally monotone and  $x \mapsto Lx - \overline{K}y$  is also maximally monotone. As the domain of this operator is  $\mathcal{H}$  and  $\partial g$  is maximally monotone,  $x \mapsto Lx - \overline{K}y + \partial g(x)$  is maximally monotone. It then follows from [2, Proposition 23.39] that  $\mathcal{F}$  is closed and convex.

(iii)(a): This follows from [2, Proposition 23.35].

(iii)(b): If  $L + L^*$  is positive definite then, for every  $x \in \mathcal{H} \setminus \{0\}$ ,

$$\langle x \mid Lx \rangle = \frac{1}{2} \langle x \mid (L + L^*)x \rangle > 0, \quad (47)$$

which shows that  $L$  is strictly monotone. Since  $\partial g$  is monotone, we deduce that  $L + \partial g$  is strictly monotone, and (iii)(a) allows us to conclude that  $T_\gamma$  has at most one fixed point.

(iii)(c): According to [2, Example 22.4(ii)], if  $g$  is strictly convex, then  $\partial g$  is strictly monotone.  $\lambda_{\min} \geq 0$  if and only if  $L$  is monotone.  $L + \partial g$  is then strictly monotone. Thus the result still follows from (iii)(a).

(iii)(d): If  $\lambda_{\min} \geq 0$ , because of the monotonicity and the continuity of  $L$ ,  $L + \partial g$  is maximally monotone. The result then follows from [2, Proposition 23.37].

(iii)(e): For every  $x \in \mathcal{H}$ ,

$$\langle x \mid Lx \rangle \geq \lambda_{\min} \|x\|^2, \quad (48)$$

which shows that  $L$  is strongly monotone. We deduce that  $L + \partial g$  is strongly monotone, and (iii)(d) allows us to conclude that  $\mathcal{F}$  is a singleton.

(iii)(f): According to [2, Example 22.4(iv)], if  $g$  is strongly convex, then  $\partial g$  is strongly monotone. Since  $L$  is monotone,  $L + \partial g$  is strongly monotone and the result follows from (iii)(d).

(iv)(a): Let  $A$  and  $B$  be defined by (24) and (25), respectively. We have thus

$$L + \partial g = A + \partial g + B. \quad (49)$$

According to Proposition 3.4,

$$\lambda_{\min} = \inf_{x \in \mathcal{H} \atop \|x\|=1} \langle x \mid Lx \rangle = \inf_{x \in \mathcal{H} \atop \|x\|=1} \langle x \mid Ax \rangle \geq 0, \quad (50)$$

which implies that  $A$  is maximally monotone. As  $\partial g$  is maximally monotone and  $\text{dom } A = \mathcal{H}$ ,  $A + \partial g$  is maximally monotone. Since  $B$  is a skewed continuous linear operator, it is also maximally monotone and  $A + \partial g + B$  is maximally monotone. According to Lemma 3.3(iii),  $A \neq 0$  and, since it is self-adjoint, it is  $1/\|A\|$ -cocoercive. It then follows from [2, Proposition 25.16] that  $A$  is  $3^*$  monotone. According to [2, Example 2.13],  $\partial g$  is  $3^*$  monotone. By invoking [2, Proposition 25.22],  $A + \partial g$  is thus  $3^*$  monotone. Since  $\text{dom } B = \mathcal{H} = \text{dom } (A + \partial g)$ , it can be deduced from the Brézis-Haraux theorem (see [2, Corollary 25.27(ii)]) that  $A + \partial g + B$  is surjective.

(iv)(b): The function defined by (45) also reads

$$h: x \mapsto \frac{1}{2} \langle x \mid Ax \rangle + g(x). \quad (51)$$

We have seen that  $A$  is self-adjoint and monotone (i.e. positive semi-definite and self-adjoint). Let  $z \in \mathcal{H}$ . Minimizing

$$x \mapsto h(x) - \langle x \mid z \rangle \quad (52)$$

is thus a convex optimization problem. A classical necessary condition for this problem to admit a solution is that  $h$  is coercive. In turn, if  $x$  is a solution to the optimization problem (51), it follows from Fermat's rule that

$$z \in \partial h(x) = Ax + \partial g(x). \quad (53)$$

Since  $z$  can be chosen arbitrarily, this shows that  $A + \partial g$  is surjective. The fact that  $\mathcal{F} \neq \emptyset$  then follows from (iv)(a).

(iv)(c)-(iv)(d): Let  $\gamma \in ]0, 2\eta]$  where  $\eta$  is the cocoercivity constant of  $L$  and let  $W = \text{Id} - \gamma L$ . According to Proposition 3.9,  $\|W\|_S \leq 1$  and  $\text{Id} - W = \gamma L$  is monotone. In addition,  $\text{dom } g^* = \mathcal{H}$  if and only if  $g$  is supercoercive [2, Proposition 14.15]. The results then follow from [18, Proposition 4.3(vi)(d)].  $\square$

By design, equation (44) shows that any fixed point of  $T_\gamma$  is a solution to an equilibrium instead of being defined from some optimality condition. In the context of Remark 3.7, the existence of a unique fixed point  $\tilde{x}$  for  $T_\gamma$  follows from the above result. This point can be viewed as an approximation to the minimizer of Problem (10) whose error is bounded in the following theorem.

**Theorem 3.11** *Assume that the following hold.*

- (i)  $L$  is cocoercive.
- (ii) Let  $\nu \in [0, +\infty[$  be the strong convexity modulus of  $g$ . Either  $\nu > 0$  or  $\lambda_{\min} \neq 0$ .
- (iii)  $\hat{x}$  is a solution to the minimization problem (10).

Then there exists a unique solution  $\tilde{x}$  to (44) and the following upper bound on the error incurred by the mismatch holds:

$$\|\tilde{x} - \hat{x}\| \leq \chi \|(H^* - \bar{K})(H\hat{x} - y)\|, \quad (54)$$

where

$$\chi = \inf_{\gamma \in ]0, 2\eta]} \frac{\gamma}{1 + \gamma\nu - \|\text{Id} - \gamma L\|_S} \leq \frac{1}{\nu + 2\lambda_{\min}}. \quad (55)$$

*Proof.* According to Proposition 3.4(i),  $\lambda_{\min} \geq 0$ .

If  $\lambda_{\min} > 0$ , according to Proposition 3.10(iii)(e), (44) has a unique solution  $\tilde{x}$ .

If  $\lambda_{\min} = 0$ , then  $\nu > 0$ , which means that  $g$  is  $\nu$ -strongly convex. It then follows from Proposition 3.10(iii)(f) that (44) has a unique solution  $\tilde{x}$ .

Let  $\gamma \in ]0, +\infty[$ . According to Proposition 3.10(i),  $\tilde{x} \in \text{Fix } T_\gamma$ , that is

$$\tilde{x} = \text{prox}_{\gamma g}((1 - \gamma\kappa)\tilde{x} - \gamma\bar{K}(H\tilde{x} - y)), \quad (56)$$

and we also know that

$$\hat{x} = \text{prox}_{\gamma g}((1 - \gamma\kappa)\hat{x} - \gamma H^*(H\hat{x} - y)). \quad (57)$$

We can write  $g = h + \nu/2\|\cdot\|^2$  where  $h \in \Gamma_0(\mathcal{H})$ , which implies that

$$(\forall x \in \mathcal{H}) \quad \text{prox}_{\gamma g}(x) = \text{prox}_{\frac{\gamma}{1+\gamma\nu}h}\left(\frac{x}{1+\gamma\nu}\right). \quad (58)$$

As  $\text{prox}_{\frac{\gamma}{1+\gamma\nu}h}$  is nonexpansive, we deduce that

$$\begin{aligned} \|\tilde{x} - \hat{x}\| &\leq \frac{1}{1 + \gamma\nu} \|(1 - \gamma\kappa)\tilde{x} - \gamma\bar{K}(H\tilde{x} - y) - \\ &\quad (1 - \gamma\kappa)\hat{x} + \gamma H^*(H\hat{x} - y)\| \\ &= \frac{1}{1 + \gamma\nu} \|(\text{Id} - \gamma L)(\tilde{x} - \hat{x}) + \gamma(H^* - \bar{K})(H\hat{x} - y)\| \\ &\leq \tau_\gamma \|\tilde{x} - \hat{x}\| + \frac{\gamma}{1 + \gamma\nu} \|(H^* - \bar{K})(H\hat{x} - y)\| \end{aligned} \quad (59)$$

with

$$\tau_\gamma = \frac{\|\text{Id} - \gamma L\|_S}{1 + \gamma\nu}. \quad (60)$$

In addition, according to Proposition 3.9, when  $\gamma \leq 2\eta$ ,

$$\|\text{Id} - \gamma L\|_S \leq 1. \quad (61)$$

This ensures that  $\tau_\gamma < 1$ , when  $\nu > 0$ .

Assume now that  $\lambda_{\min} > 0$ . If  $\gamma < 2\eta$ , (39) yields

$$\|\text{Id} - \gamma L\|_S < 1, \quad (62)$$

which also guarantees that  $\tau_\gamma < 1$ .

In summary, if  $\gamma < 2\eta$ , it can be deduced from (59) that

$$\|\tilde{x} - \hat{x}\| \leq \frac{\gamma}{(1 - \tau_\gamma)(1 + \gamma\nu)} \|(H^* - \bar{K})(H\hat{x} - y)\|, \quad (63)$$

which leads to (54). In addition, according to (60) and (39),

$$(1 - \tau_\gamma)(1 + \gamma\nu) \geq \gamma \left( \nu + \left(2 - \frac{\gamma}{\eta}\right) \lambda_{\min} \right) > 0. \quad (64)$$

By noticing that

$$\sup_{\gamma \in ]0, 2\eta[} \nu + \left(2 - \frac{\gamma}{\eta}\right) \lambda_{\min} = \nu + 2\lambda_{\min}, \quad (65)$$

the upper bound on  $\chi$  is obtained.  $\square$

### Remark 3.12

(i) Under the assumptions of the above proposition, we deduce from (54) that

$$\|\tilde{x} - \hat{x}\| \leq \chi \|H^* - \bar{K}\|_S \|H\hat{x} - y\|. \quad (66)$$

This upper bound tells us that the error depends on the data error (which encompasses noise and modelling errors) and the norm of the mismatch on the adjoint.

(ii) In addition, the parameter  $\chi$  depends on the strong convexity modulus  $\nu$  and on the quadratic regularization parameter  $\kappa$ . Indeed, the larger  $\kappa$ , the larger  $\lambda_{\min}$ . The upper bound in (55) shows that increasing  $\nu$  or  $\kappa$  allows us to decrease the distance to the true minimizer  $\hat{x}$ . At the same time, these parameters control the regularization term in (10) so that large values of them can introduce a bias in the recovery of the true signal. One should therefore seek values of these parameters balancing these two effects.

**Remark 3.13** It follows from [24, Theorem 3.3] that, when  $g = 0$  and  $H^*H + \kappa\text{Id}$  is invertible,

$$\|\tilde{x} - \hat{x}\| \leq \frac{1}{\kappa} \|(H^* - \bar{K})(H\hat{x} - y)\| + o(\|H^* - \bar{K}\|). \quad (67)$$

This bound is less tight than the one in (54)-(55) if  $2\lambda_{\min} > \kappa \Leftrightarrow \kappa > -2\tilde{\lambda}_{\min}$ , where  $\tilde{\lambda}_{\min}$  is the minimum spectral value of  $(\bar{K}H + H^*\bar{K}^*)/2$ .

## 3.4. Convergence result

We present a first result concerning the averageness properties of operator  $T_\gamma$  with  $\gamma \in ]0, +\infty[$ .

**Lemma 3.14** *Let  $\eta \in ]0, +\infty[$ , let  $\gamma \in ]0, 2\eta[$ , let*

$$\bar{\alpha} = \frac{1}{2 - \gamma/(2\eta)} \in \left] \frac{1}{2}, 1 \right[ , \quad (68)$$

*and let  $\bar{W} = \text{Id} - \gamma L$ . The following properties hold.*

$$L \text{ is } \eta\text{-cocoercive} \quad (69)$$

$$\Leftrightarrow \bar{W} \text{ is } \gamma/(2\eta)\text{-averaged} \quad (70)$$

$$\Rightarrow (\forall x \in \mathcal{H}) \quad \|\bar{W}x - 2(1 - \bar{\alpha})x\| + \|\bar{W}x\| \leq 2\bar{\alpha}\|x\| \quad (71)$$

$$\Rightarrow T_\gamma \text{ is } \bar{\alpha}\text{-averaged}. \quad (72)$$

*Proof.* If  $\gamma < 2\eta$  and  $L$  is  $\eta$ -cocoercive, then the first equivalence holds [2, Proposition 4.39].

Let us now show that (70) implies (71). Set  $\alpha = \gamma/(2\eta)$ . Since  $\bar{W}$  is  $\alpha$ -averaged, there exists a nonexpansive operator  $Q: \mathcal{H} \rightarrow \mathcal{H}$  such that  $\bar{W} = (1 - \alpha)\text{Id} + \alpha Q$ . We have then, for every  $x \in \mathcal{H}$ ,

$$\begin{aligned} & \|\bar{W}x - 2(1 - \bar{\alpha})x\| + \|\bar{W}x\| \\ &= \|(1 - \alpha)x + \alpha Qx - 2(1 - (2 - \alpha)^{-1})x\| + \|(1 - \alpha)x + \alpha Qx\| \\ &= \alpha\|Qx - (1 - \alpha)(2 - \alpha)^{-1}x\| + \|(1 - \alpha)x + \alpha Qx\| \\ &= \alpha \sqrt{\|Qx\|^2 + \left(\frac{1 - \alpha}{2 - \alpha}\right)^2 \|x\|^2 - 2\frac{1 - \alpha}{2 - \alpha} \langle x | Qx \rangle} \\ & \quad + \sqrt{\alpha^2 \|Qx\|^2 + (1 - \alpha)^2 \|x\|^2 + 2\alpha(1 - \alpha) \langle x | Qx \rangle} \\ &\leq \varphi(\theta)\|x\|, \end{aligned} \quad (73)$$

where

$$\varphi(\theta) = \alpha \sqrt{1 + \left(\frac{1 - \alpha}{2 - \alpha}\right)^2 - 2\frac{1 - \alpha}{2 - \alpha} \theta} + \sqrt{\alpha^2 + (1 - \alpha)^2 + 2\alpha(1 - \alpha)\theta}.$$

In the last inequality, we have set  $\langle x | Qx \rangle = \theta\|x\|\|Qx\|$  and used the nonexpansiveness of  $Q$ . Let us now study function  $\varphi$  on  $[-1, 1]$ . The derivative  $\varphi'$  of this function is such that

$$\begin{aligned} \alpha^{-1}(1 - \alpha)^{-1}\varphi'(\theta) &= \frac{1}{\sqrt{\alpha^2 + (1 - \alpha)^2 + 2\alpha(1 - \alpha)\theta}} \\ & \quad - \frac{1}{\sqrt{(2 - \alpha)^2 + (1 - \alpha)^2 - 2(1 - \alpha)(2 - \alpha)\theta}}. \end{aligned}$$

Therefore,  $\varphi'(\theta) \geq 0$  since

$$\begin{aligned} \alpha^2 + (1 - \alpha)^2 + 2\alpha(1 - \alpha)\theta &\leq (2 - \alpha)^2 + (1 - \alpha)^2 - 2(1 - \alpha)(2 - \alpha)\theta \\ \Leftrightarrow \theta &\leq 1, \end{aligned} \quad (74)$$

where we used the fact that  $\alpha \in ]0, 1[$ . This shows that  $\varphi$  is increasing on  $[-1, 1]$  and we deduce from (73) that

$$\|\bar{W}x - 2(1 - \bar{\alpha})x\| + \|\bar{W}x\| \leq \varphi(1)\|x\| = \frac{2}{2 - \alpha}\|x\| = 2\bar{\alpha}\|x\|. \quad (75)$$

Since  $\text{prox}_{\gamma f}$  is firmly nonexpansive, it follows from [18, Theorem 3.8], that when (71) holds,  $T_\gamma$  is  $\bar{\alpha}$ -averaged.  $\square$

Given the above properties, the convergence of the mismatched PGA is guaranteed by the following result.

**Proposition 3.15** *Assume that  $L$  is  $\eta$ -cocoercive with  $\eta \in ]0, +\infty[$ . Let  $\gamma \in ]0, 2\eta[$  and  $\delta = 2 - \gamma/(2\eta)$ . Let  $(\theta_n)_{n \in \mathbb{N}}$  be a sequence in  $[0, \delta]$  such that  $\sum_{n \in \mathbb{N}} \theta_n(\delta - \theta_n) = +\infty$ . Suppose that  $\mathcal{F} \neq \emptyset$ . Then the sequence  $(x_n)_{n \in \mathbb{N}}$  generated by Algorithm (12) converges weakly to a point  $\tilde{x} \in \mathcal{F}$ . In addition, if  $\lambda_{\min} \neq 0$  and, for every  $n \in \mathbb{N}$ ,  $\omega_n = 0$  and  $\theta_n \in [\underline{\theta}, 1]$  with  $\underline{\theta} \in ]0, +\infty[$ , then  $(x_n)_{n \in \mathbb{N}}$  converges linearly.*

*Proof:* For every  $n \in \mathbb{N}$ , let  $W_n = (1 - \gamma\kappa)\text{Id} - \gamma K_n H$ , let  $b_n = \gamma K_n y$ , let  $\bar{W} = \text{Id} - \gamma L$ , and let  $\bar{b} = \gamma \bar{K} y$ . Then (12) reads, for every  $n \in \mathbb{N}$ ,

$$x_{n+1} = x_n + \theta_n(\text{prox}_{\gamma g}(W_n x_n + b_n) - x_n). \quad (76)$$

The algorithm can thus be interpreted as an instance of the recurrent neural network investigated in [18] with  $m = 1$  layer. It follows from Lemma 3.14 that [18, Condition 3.1] holds.

In addition, as a consequence of Assumption 3.1, [18, Assumption 5.1] is satisfied since

$$\sum_{n \in \mathbb{N}} \|W_n - \bar{W}\|_S \leq \gamma \|H\|_S \sum_{n \in \mathbb{N}} \omega_n < +\infty \quad (77)$$

$$\sum_{n \in \mathbb{N}} \|b_n - \bar{b}\| \leq \gamma \|y\| \sum_{n \in \mathbb{N}} \omega_n < +\infty. \quad (78)$$

The convergence of  $(x_n)_{n \in \mathbb{N}}$  to  $\tilde{x} \in \mathcal{F}$  can then be deduced from [18, Theorem 5.4].

Assume now that  $\lambda_{\min} \neq 0$ ,  $(\forall n \in \mathbb{N}) \omega_n \equiv 0$  and  $\theta_n \in [\underline{\theta}, 1]$ . It follows from (12) and (56) that, for every  $n \in \mathbb{N}$ ,

$$\begin{aligned} x_{n+1} - \tilde{x} &= (1 - \theta_n)(x_n - \tilde{x}) \\ &+ \theta_n(\text{prox}_{\gamma g}((\text{Id} - \gamma L)x_n + \gamma \bar{K} y) - \text{prox}_{\gamma g}((\text{Id} - \gamma L)\tilde{x} + \gamma \bar{K} y)). \end{aligned} \quad (79)$$

Using the nonexpansivity of the proximity operator and the triangle inequality yield

$$\begin{aligned} \|x_{n+1} - \tilde{x}\| &\leq (1 - \theta_n)\|x_n - \tilde{x}\| + \theta_n\|(\text{Id} - \gamma L)(x_n - \tilde{x})\| \\ &\leq (1 - \theta_n + \theta_n\|\text{Id} - \gamma L\|_S)\|x_n - \tilde{x}\|. \end{aligned} \quad (80)$$

By using now Proposition 3.9, we deduce that

$$\begin{aligned} \|x_{n+1} - \tilde{x}\| &\leq \left(1 - \theta_n + \theta_n \sqrt{1 - \gamma \left(2 - \frac{\gamma}{\eta}\right) \lambda_{\min}}\right) \|x_n - \tilde{x}\| \\ &\leq \rho \|x_n - \tilde{x}\|, \end{aligned} \quad (81)$$

where

$$\rho = 1 - \left(1 - \sqrt{1 - \gamma \left(2 - \frac{\gamma}{\eta}\right) \lambda_{\min}}\right) \underline{\theta} \in ]0, 1[. \quad (82)$$

We deduce that, for every  $n \in \mathbb{N}$ ,  $\|x_n - \tilde{x}\| \leq \rho^n \|x_0 - \tilde{x}\|$ , which shows the linear convergence of  $(x_n)_{n \in \mathbb{N}}$ .  $\square$

We now see that the cocoercivity constant of  $L$  is useful to obtain an upper bound on the gradient descent parameter.

**Remark 3.16** If  $L$  is self-adjoint positive (i.e.  $\beta = 0$  and  $\lambda_{\min} \geq 0$ ), then it follows from Proposition 3.4 that  $L$  is  $\eta$ -cocoercive with  $1/\eta = \lambda_{\max} = \|L\|_{\text{S}}$ . Proposition 3.15 thus leads to  $2/\|L\|_{\text{S}}$  as a strict upper bound on step-size  $\gamma$  in order to guarantee the convergence of the algorithm. This allows us to recover the classical upper bound on the step-size for Algorithm (12) in the special case when  $\bar{K} = H^*$ .

**Remark 3.17** When  $g = 0$ ,  $\theta_n \equiv 1$ , and  $\mathcal{H} = \mathbb{R}^N$ , (12) becomes a linear recursive equation and tools from matrix analysis can be employed to derive the following necessary and sufficient convergence conditions [24, Theorem 3.1]:

$$(\forall j \in \mathbb{J}) \quad \gamma < 2 \frac{\operatorname{Re} \zeta_j}{|\zeta_j|^2} \quad (83)$$

$$\operatorname{Re} \zeta_j > 0, \quad (84)$$

where  $(\zeta_j)_{j \in \mathbb{J}}$  are the nonzero eigenvalues of  $L$ . It is easy to show that, for every  $j \in \mathbb{J}$ ,  $\lambda_{\min} \leq \operatorname{Re} \zeta_j$ . Therefore, if  $\lambda_{\min} > 0$ , (84) is satisfied. Then, it follows from Propositions 3.4(ii) and Corollary 3.6(i) that a sufficient and necessary condition for  $L$  to be  $\eta$ -cocoercive is  $\eta \leq \bar{\eta}$  where  $\bar{\eta}$  is given by (38). Since Proposition 3.15 guarantees the convergence of (12) when  $\gamma \in ]0, 2\bar{\eta}[$ , we deduce that

$$(\forall j \in \mathbb{J}) \quad \bar{\eta} \leq \frac{\operatorname{Re} \zeta_j}{|\zeta_j|^2}. \quad (85)$$

This emphasizes that, in the presence of adjoint mismatch, the cocoercivity of  $L$  only provides a sufficient condition for the convergence of PGA.

#### 4. Numerical experiments

As stated in the introduction, the proposed method is applicable to a wide range of inverse problems. In this section, the focus is placed on two X-ray 2D tomographic image reconstruction problems that appear in image-guidance for interventional radiology and surgery.

We aim at recovering an image  $\bar{x}$  having  $N$  pixels in the Euclidean space  $\mathcal{H} = \mathbb{R}^N$  from a set of tomographic measurements  $y \in \mathcal{G} = \mathbb{R}^M$ . In the observation model (1),  $H \in \mathbb{R}^{M \times N}$  is the projection matrix and  $b$  is an additive i.i.d. zero-mean Gaussian noise with standard deviation  $\sigma \in ]0, +\infty[$ . We consider the frequent problem arising in this imaging modality, where a constant error is introduced on the adjoint, when implementing the PGA algorithm, so that  $K_n \equiv \bar{K}$ . To quantify the error introduced by  $\bar{K} \in \mathbb{R}^{N \times M}$ , we define  $\delta$  as the average over 20 realizations of the ratio  $\langle Hu \mid v \rangle / \langle u \mid \bar{K}v \rangle$  with  $(u, v)$  i.i.d. uniformly sampled on  $([0, 1]^N)^2$ . The farther  $\delta$  from 1, the farther  $\bar{K}$  from  $H^*$ . In addition to this coupling ratio, we provide a measure of asymmetry  $\mu$  defined as  $\|KH - H^*K^*\|_{\text{F}} / (2\|KH\|_{\text{F}})$  where  $\|\cdot\|_{\text{F}}$  denotes the Frobenius norm.

In our first experiment, a geometric abdomen phantom is reconstructed with a truncated field of view. This corresponds to the common setup where the detector is not large enough to model large body parts such as the abdomen. In the second experiment, we perform a joint reconstruction and segmentation of the metal part in a region of interest (ROI) of another geometric phantom from undersampled projections. For such class of under-determined problems, iterative reconstruction methods have proven their superiority over filtered backprojection [45]. Note that, in both experiments, the noise variance  $\sigma^2$  is kept small to emphasize that sub-sampling



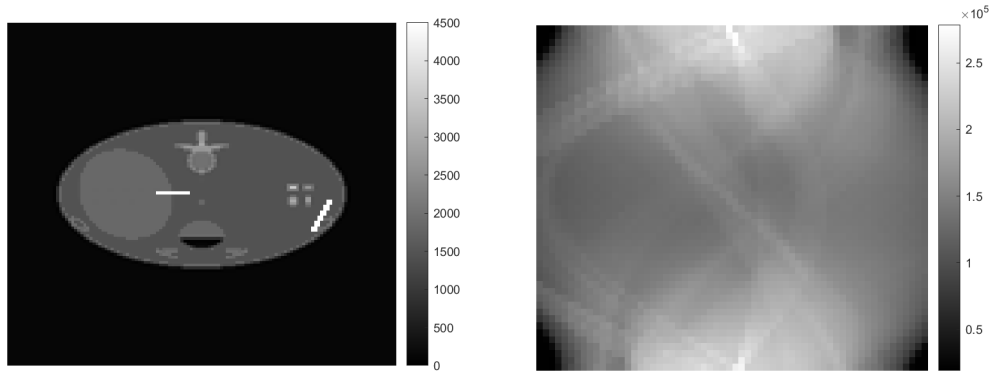
is the main issue (as it is the case in interventional C-arm CBCT imaging) while still avoiding to commit an inverse crime.

All the simulations were performed with the ASTRA Toolbox [47] implemented on Matlab, which allows the explicit computation of matrices  $H$  and  $\bar{K}$ . In both experiments, matrix  $H$  was computed using ASTRA line-length ray driven projector [51], while  $\bar{K}$  was generated using the pixel-driven backprojector [48]. Note that the latter corresponds to the default choice in the GPU implementation for backprojection in ASTRA.

#### 4.1. Example 1: Reconstruction of a geometric abdomen from undersampled projections

**4.1.1. Problem statement:** We simulated a scan of an abdomen of size 45 cm based on a modified version of the forbild abdomen phantom. It is made of a vertebrae set to 3000, high intensity inserts ranging from 4000 to 4500, a liver area set to 1840. These values correspond to positive Hounsfields Unit (HU) such that air is 0 HU and water is 1000 HU. The source-to-object distance and the source-to-image distance were respectively set to 800 mm and 1200 mm leading to a magnification factor of 1.5 as can be found on clinical scanners. The associated sinogram was computed in fan beam geometry over  $180^\circ$  using 50 regularly spaced angular steps. The projection and backprojection operators were rescaled by  $\pi/50$  so as to match with the analytical definition of the backprojector [31] and to make the parametrization independent from the number of projections. The detector has 62 bins of size 6.4 mm, so that  $M = 62 \times 50$ , and the image is reconstructed on a discrete grid of  $N = 128 \times 128$  pixels, with size  $1.5 \times 6.4 = 4.26$  mm. The image reconstruction problem is undetermined, due to the small detector field of view (FOV), the limited angular coverage and the low density of the object. Noise standard deviation  $\sigma$  is chosen such that  $\|b\|/\|H\bar{x}\| \approx 6.3 \times 10^{-5}$ . Figure 1 shows the phantom  $\bar{x}$  and the data  $y$ .

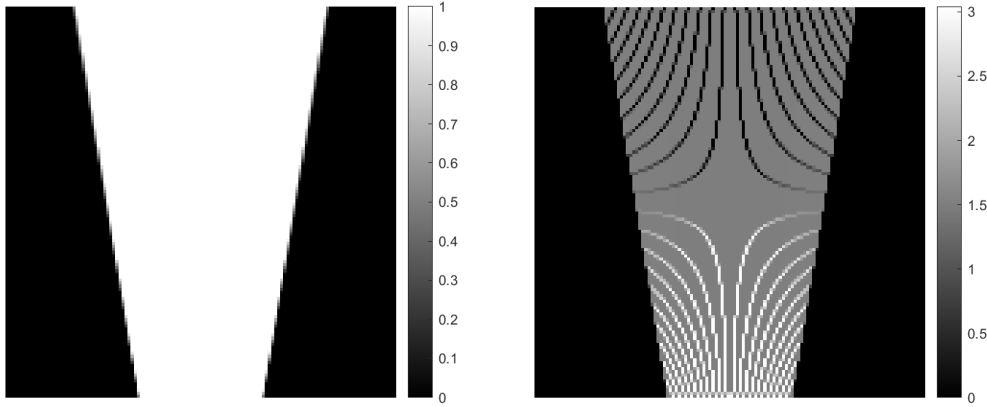
With those settings,  $H^*$  contains 1.08% nonzero elements whereas this proportion



**Figure 1.** Phantom  $\bar{x}$  (left) and sinogram  $y$  (right)

decreases to 0.89% for  $\bar{K}$ . The coupling ratio is  $\delta = 1.151$  and the asymmetry metric  $\mu$  is equal to 0.159. Figure 2 shows the backprojection of constant measurements at a single angle using either  $\bar{K}$  or  $H^*$ . A high frequency Moire pattern is visible when using  $H^*$  (right image) due to the redundancy introduced by oversampling the

projection, while the backprojected view remains uniform with  $\overline{K}$  (left image).



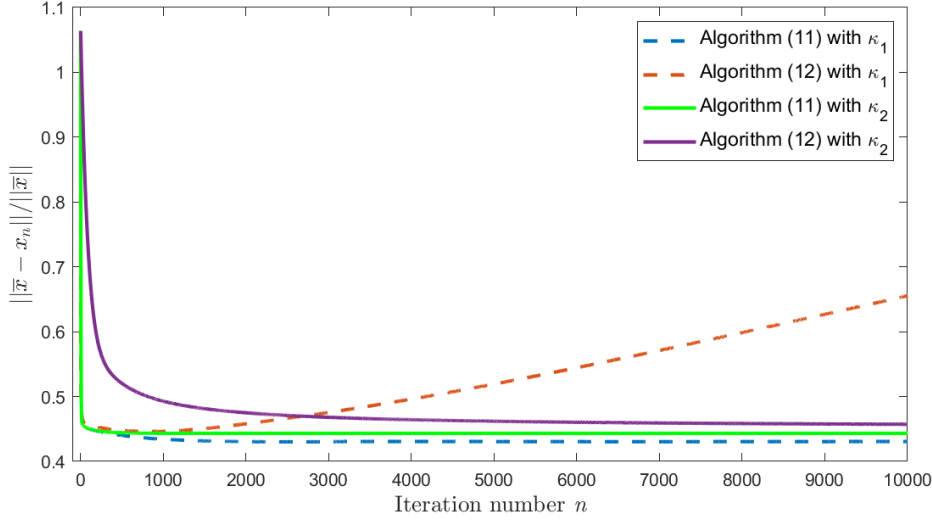
**Figure 2.** Backprojection of a uniform view with  $\overline{K}$  (left) and  $H^*$  (right)

An estimate of  $\bar{x}$  is obtained by adopting a sparse inducing formulation, reminiscent from the literature on compressive sensing [9,25,26,30,36]. We solve the penalized least squares problem (10) with  $g = \rho \|W \cdot\|_1$ ,  $W \in \mathbb{R}^{N \times N}$  being the orthogonal Symlet 2 wavelet transform on 2 resolution levels, and  $\rho > 0$  the associated regularization parameter.

We ran Algorithms (11) and (12), for two settings  $\kappa_1$  and  $\kappa_2$  of parameter  $\kappa$  such that  $L$  is not cocoercive with  $\kappa_1$ , but becomes cocoercive with  $\kappa_2$ . In such case, the condition given in Proposition 3.10(iii)(e) holds, which proves the existence of a unique fixed point of scheme (12) and its convergence is ensured according to Proposition 3.15. We set  $\kappa_1 = 10^{-2}$ . Moreover, following Remark 3.7,  $\kappa_2$  is set as  $\tilde{\lambda}_{\min} + 10^{-2}$ . The eigenvalue  $\tilde{\lambda}_{\min} = 1.61$  is computed using the Matlab function `eigs`. Note that despite the fact that matrices  $H$  and  $\overline{K}$  were stored in these experiments, matrix-free iterative methods can be used to compute the dominant and the smallest eigenvalues of operator  $(L+L^*)/2$  thus complying with practical implementations of the projector and backprojector for higher dimensional problems. Moreover, to bypass the need for the exact adjoints of  $H$  and  $\overline{K}^*$  while computing minimum eigenvalues, we refer to the strategy in [24]. We set additionally the regularization hyperparameter  $\rho$  to 600, after tuning with a grid search so as to reach satisfying visual results, and the relaxation parameter  $\theta_n \equiv 1$ . For the coupled settings  $(H^*, \kappa_1)$ ,  $(\overline{K}, \kappa_1)$  and  $(H^*, \kappa_2)$ , stepsize  $\gamma$  was set to  $1.9/(\|H\|_S^2 + \kappa) = 2.9 \times 10^{-3}$ . For  $(\overline{K}, \kappa_2)$ ,  $\gamma$  is chosen equal to  $1.82 \times 10^{-5}$  in accordance with Corollary 3.6 and Proposition 3.15. The algorithms are ran until a stopping precision on the relative distance between two consecutive iterates is below  $10^{-7}$  or a maximum number of iterations of  $10^4$  is reached.

**4.1.2. Results:** Figure 3 displays the normalized mean square error (NMSE) defined as  $(\|\bar{x} - x_n\|/\|\bar{x}\|)_n$ , computed along the iterations when applying Algorithms (11) and (12). The plots confirm that, with the value  $\kappa_1$ , PGA converges when the exact adjoint  $H^*$  is used but diverges when replaced by  $\overline{K}$ , as was expected from our theoretical analysis. In the latter case, Algorithm (12) shows an initial convergence trend that reaches a minimum discrepancy point close to the minimizer obtained with  $H^*$  before diverging. For value  $\kappa_2$ , both Algorithms (11) and (12) converge to fixed points that

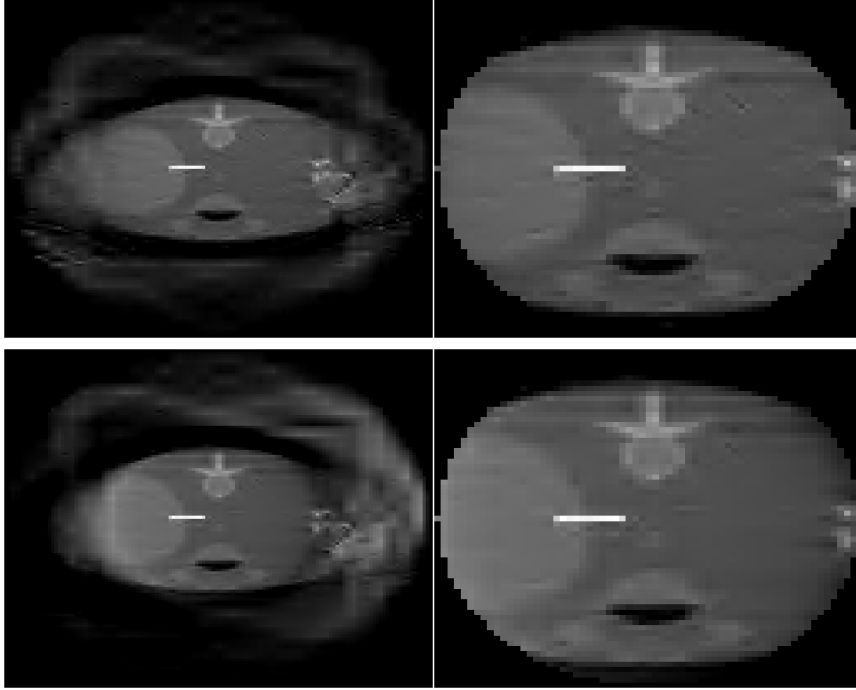
are close to each other, again confirming our theoretical analysis. The corresponding NMSE values are 0.4432 and 0.4572, respectively. PGA without mismatch requires less iterations to reach convergence than its perturbed version using  $\bar{K}$ . Note that, in a real context, practitioners often use early stopping to avoid the potential negative effects of the adjoint mismatch. Nevertheless it is difficult for the user to know when the iterations should be stopped so as to reach this intermediary good solution, hence the result is often suboptimal. Our analysis shows that one can still use an inaccurate adjoint without resorting to such an empirical rule.



**Figure 3.** Decay of the error along iterations for Algorithms (11) and (12) and two choices of  $\kappa$  parameter.

Reconstruction results are displayed in Figures 4 and 5 using the same windowing. Let us remark that, due to the use of a short detector, the projections suffer from axial truncation. The set of pixels of the image whose projections belong to then define the so-called image FOV. We added a comparison with two reconstructions obtained from the standard filtered back-projection (FBP) approach in Figure 6. On the left, the image is obtained from standard FBP by zero-padding the sinogram, while on the right, the image is obtained by replicating the borders of the sinogram [35]. Only the image FOV is depicted here, since the FBP reconstruction outside this zone is not relevant. We also indicated the NMSE and the maximum absolute error (MAE), defined as  $\max_{i \in \{1, \dots, N\}} |\bar{x}_i - x_i|$ , for all the reconstructed images when compared with the ground truth. Both FBP reconstructions suffer from various artifacts (peripheral brightband artifacts, cupping, over-estimation of the values as shown in Figure 7) [41–43], in contrast with the solutions provided by our regularized iterative approach. Furthermore, when parameter  $\kappa_2$  is used, the reconstructed image obtained by PGA with the mismatched adjoint  $\bar{K}$  is very similar to the image obtained without mismatch. In contrast, combining the setting  $\kappa_1$  with the mismatched adjoint in PGA yields a reconstruction that is deteriorated by artifacts propagating from the exterior of the FOV and a higher NMSE compared to the solution obtained when using the exact  $H^*$  as shown on the reconstructed image in Figure 4 (bottom left) and the FOV

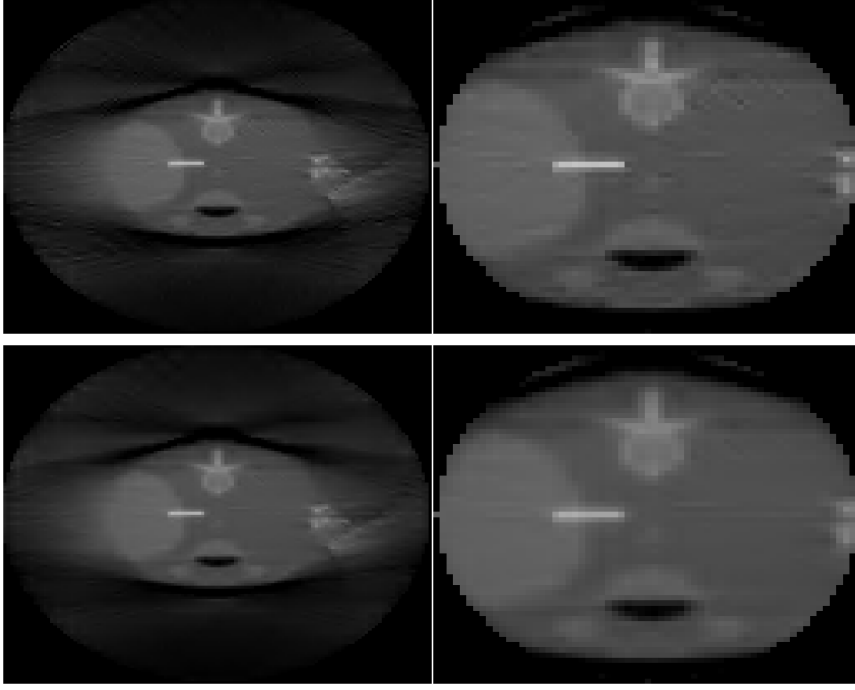
error map in Figure 8 (top right). In a nutshell, as soon as the convergence of PGA is ensured, an unmatched projector/backprojector pair gives a similar reconstruction quality than the matched pair but may lead to a slower convergence. Let us emphasize that, in practice, the decrease of the convergence rate in terms of iterations could be compensated by a reduced computation cost for operator  $\bar{K}$ . Finally, note that computing the infimum in (55) in Theorem 3.11 with a grid search gives an upper bound of  $3.25 \times 10^5$ . The actual error is  $1.0934 \times 10^4$  which is indeed lower than this upper bound, as expected from our theoretical analysis.



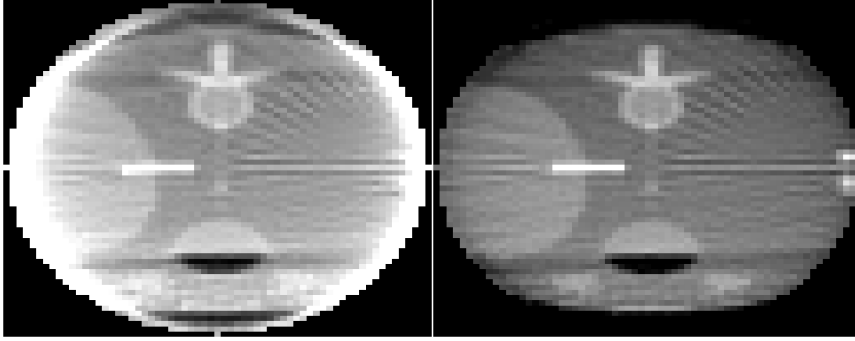
**Figure 4.** Reconstructions (left) and zoomed versions within the FOV (right) obtained using  $\kappa_1$  and either Algorithm (11), NMSE = 0.1207, MAE = 2330 (top) or Algorithm (12), NMSE = 0.1610, MAE = 3141 (bottom).

#### 4.2. Example 2: Joint object-background decomposition and super resolution reconstruction

**4.2.1. Problem statement:** In this example, we focus on a joint super resolution reconstruction and decomposition task. Flat panel detectors commonly sample projections with small pixels but at a slow frame rate, so that the angular sampling is comparatively poor. Then reconstructing the entire object on a fine grid is time consuming and produces large volumes that are also difficult to manipulate. We thus look at reconstructing a relevant ROI only, as is the case when the clinical goal is to assess the precise position of metallic needles within a soft-tissue background. With metallic device, subsampling artifacts have a minor impact on contrasted soft tissue background. A priori knowledge about the device (e.g. sparsity, high contrast, and direction [5,34]) can be used, given that the object is separated from the background.



**Figure 5.** Reconstructions (left) and zoomed versions within the FOV (right) obtained using  $\kappa_2$  and either Algorithm (11), NMSE = 0.16, MAE = 2205 (top) or Algorithm (12), NMSE = 0.1534, MAE = 2399 (bottom).

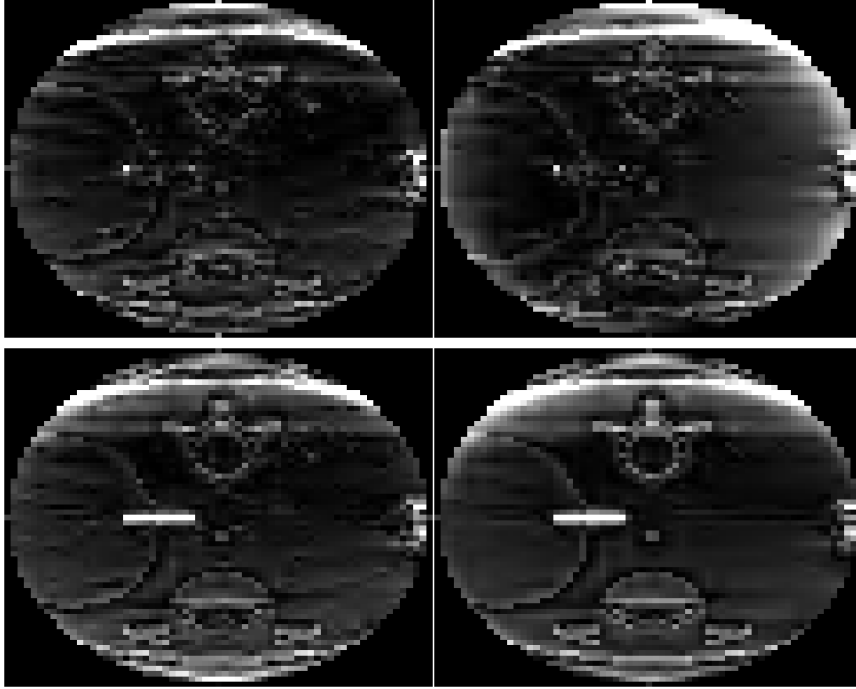


**Figure 6.** FBP reconstructions, in the FOV, with zero-padded FBP, NMSE = 1.776, MAE = 8534 (left) and extrapolated FBP by replicating the borders of the sinogram, NMSE = 0.366, MAE = 1871 (right).

The phantom grid is of  $256 \times 256$  pixels of size 0.53 mm of which the ROI, denoted by  $\bar{x}_r$ , is a patch of size  $88 \times 88$ . The simulated phantom  $\bar{x}$  and the ROI are displayed in Figure 9. The phantom projection is computed for a detector of 500 bins of 0.4 mm. The detector bins are sampled on a twice thinner grid than the pixels. The number of uniformly spaced angular positions is set to 100 only over interval  $[0^\circ, 180^\circ]$  leading to  $M = 100 \times 500$ . The source-to-object distance and the source-to-image distance were set as in our first experiment. The operators are rescaled by  $\pi/100$ . The noise



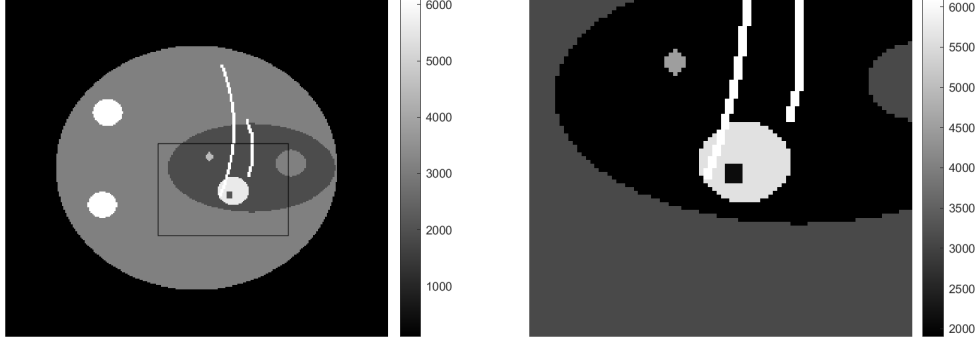
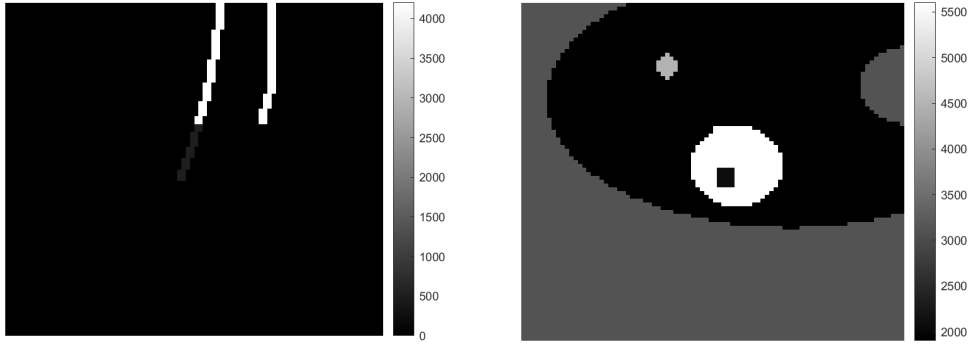
**Figure 7.** Absolute difference between the reconstructed image from FBP using replicated sinogram borders and the ground truth, within the FOV.



**Figure 8.** Absolute difference between the reconstructed image and the ground truth, within the FOV, using  $\kappa_1$  (top) or  $\kappa_2$  (bottom), and either Algorithm (11) (left) or Algorithm (12) (right).

standard deviation is such that  $\|b\|/\|H\bar{x}\| \approx 3.32 \times 10^{-5}$ .

The acquired projections contain information regarding pixels outside the ROI. In order to reduce reconstruction artifacts, we define a larger reconstruction grid, with size  $N = 140 \times 140$  containing the ROI. Let us introduce the sampling operator  $S \in \mathbb{R}^{88^2 \times 140^2}$ , which selects the ROI region within this extended image. We then aim at decomposing the spacial contents  $\bar{x}_r$  within this ROI into two maps  $S\bar{x}_m \in \mathbb{R}^{88 \times 88}$  and  $S\bar{x}_b \in \mathbb{R}^{88 \times 88}$  which describe respectively the metal component of the ROI (needles) and the tissues of the ROI as shown in Figure 10, so that  $\bar{x}_r = S\bar{x}_m + S\bar{x}_b$ .

Figure 9. Phantom  $\bar{x}$  (left) and ROI  $\bar{x}_r$  (right).Figure 10.  $S\bar{x}_m$  (left) and  $S\bar{x}_b$  (right).

Estimates of the two maps  $(\bar{x}_m, \bar{x}_b) \in \mathbb{R}^{2N}$  on the extended grid of size  $N$ , are obtained by solving the following penalized least-squares problem:

$$\begin{aligned} \underset{(x_m, x_b) \in \mathbb{R}^{2N}}{\text{minimize}} \quad & \frac{1}{2} \|y - H_r(x_m + x_b)\|_2^2 \\ & + g(x_m, x_b) + \frac{\kappa}{2} (\|x_m\|_2^2 + \|x_b\|_2^2). \end{aligned} \quad (86)$$

Furthermore, we define  $g(x_m, x_b) = \rho \text{DTV}_{a,\theta}(x_m) + \beta \text{TV}(x_b) + \alpha \|x_m\|_1 + \iota_{[0,+\infty[^N]}(x_m) + \iota_{[0,+\infty[^N]}(x_b)$  where  $\iota_C$  denotes the indicator function of set  $C$  and  $(\rho, \beta, \alpha, \kappa) \in [0, +\infty[^4$ . Hereabove, the TV term, acting on the background image, stands for the total variation regularization [46], defined as

$$(\forall u \in \mathbb{R}^N) \quad \text{TV}(u) = \sum_{i=1}^N \sqrt{(\Delta_i^h u)^2 + (\Delta_i^v u)^2} \quad (87)$$

with  $(\Delta_i^h)^* \in \mathbb{R}^N$ ,  $(\Delta_i^v)^* \in \mathbb{R}^N$ , the horizontal and vertical discrete gradient operators at location  $i$  (assuming zero-padding), respectively. Furthermore, given that the sought map  $\bar{x}_m$  is sparse, and contains needles following about the same direction, we use both an  $\ell_1$  penalty and the directional total variation introduced in [6], defined, for every  $u \in \mathbb{R}^N$ , as  $\text{DTV}_{a,\theta}(u) = \sum_{i=1}^N \|D_{i,a,\theta} u\|$  where  $D_{i,a,\theta} \in \mathbb{R}^{2 \times N}$  allows to

compute the two directional derivatives at the pixel  $i$ , parametrized by an angular direction  $\theta \in [0, 180^\circ[$ , and a scaling factor  $a > 0$ , i.e.

$$D_{i,a,\theta}u = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \Delta_i^h u \\ \Delta_i^v u \end{pmatrix}. \quad (88)$$

Let  $H = (H_r, H_r)$  in  $\mathbb{R}^{M \times 2N}$ . Equation (86) can be rewritten as

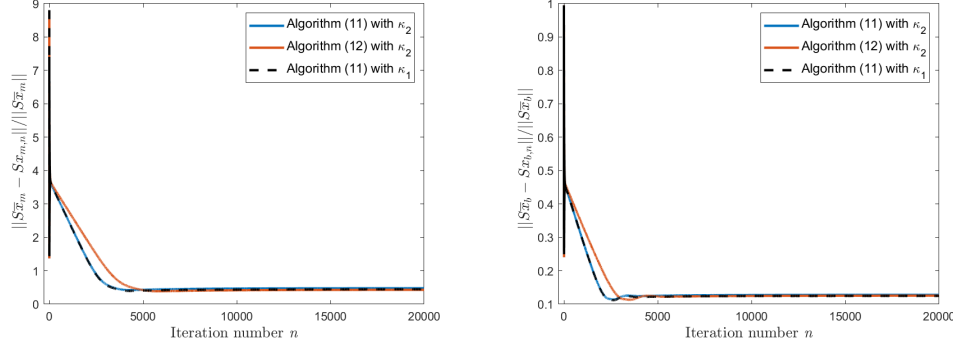
$$\underset{z=(x_m^\top, x_b^\top)^\top \in \mathbb{R}^{2N}}{\text{minimize}} \quad \frac{1}{2} \|y - Hz\|_2^2 + h(z) + \frac{\kappa}{2} \|z\|_2^2 + i_{[0,+\infty[2N]}(z), \quad (89)$$

with  $h: z = (x_m^\top, x_b^\top)^\top \mapsto \rho DTV_{a,\theta}(x_m) + \alpha \|x_m\|_1 + \beta TV(x_b)$ .

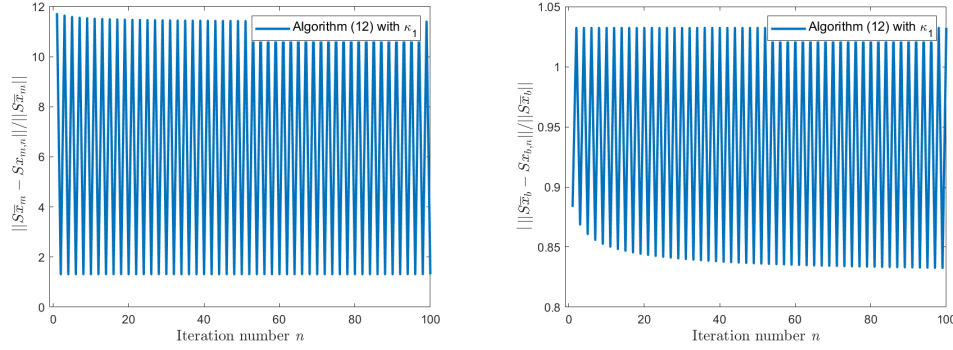
The coupling ratio between  $H$  and its associated adjoint approximation  $\bar{K} = (\bar{K}_r, \bar{K}_r)$  is  $\delta = 0.75$  and the asymmetry metric is  $\mu = 0.0418$ . The proximity operator of  $h$  does not have a closed form, hence it will be approximated by using inner iterations of the dual forward-backward algorithm [14, 15] with a stopping precision of  $10^{-8}$ . We wet  $\theta_n \equiv 1$ ,  $\rho = 5500$ ,  $\beta = 2950$ ,  $\alpha = 500$ ,  $a = 0.2$ , and  $\theta = 10^\circ$ . As in our first experiment, two values of  $\kappa$  are tested, namely  $\kappa_1 = 10^{-2}$  and  $\kappa_2 = 0.2438$ .  $L$  is guaranteed to be cocoercive for  $\kappa = \kappa_2$ , but not for  $\kappa = \kappa_1$ . Here again the existence and uniqueness of the fixed point of scheme (12) are guaranteed for  $\kappa = \kappa_2$  because the condition in Proposition 3.10(iii)(e) is fulfilled. Furthermore, for the settings  $(H^*, \kappa_1)$ ,  $(\bar{K}, \kappa_1)$  and  $(H^*, \kappa_2)$ , the stepsize  $\gamma$  is set respectively to  $2 \times 10^{-3}$  while for  $(\bar{K}, \kappa_2)$ ,  $\gamma$  is set to  $1.5 \times 10^{-3}$ . The stopping precision on the relative distance between two consecutive iterates is  $10^{-7}$  and the maximum number of iterations is  $2 \times 10^4$ .

**4.2.2. Results:** In Figures 11 and 12, we plot the relative errors between the ground truth metal map  $S\bar{x}_m$  and tissues map  $S\bar{x}_b$ , cropped to the ROI region, and their estimate along the iterations. In Figure 12, one sees that the iterates obtained from (12) with  $\kappa_1$  are unstable. Oscillations hamper the convergence of scheme (12). The stopping convergence criterion is never met and at the end of the  $2 \times 10^4$  iterations, the maps cannot be reconstructed. Figure 11 shows that for the three other cases, the algorithm stops in a phase where the errors associated with both maps are simultaneously decreasing. These plots confirm that, with the setting  $\kappa_1$ , only Algorithm (11) (i.e. PGA without adjoint mismatch) converges. For  $\kappa_2$ , Algorithms (11) and (12) converge to two fixed points that are quite close to each other and to the exact solution. We also report the value of the objective function with  $\kappa_2$  over the iterations. As expected, at convergence, the value of the objective function is higher with a mismatch than without a mismatch. Figure 13 shows the reconstructed maps within the ROI region, obtained with Algorithm (11) and  $\kappa_1$ , Algorithm (11) and  $\kappa_2$ , Algorithm (12) and  $\kappa_2$ . Upon visual inspection, the two restored components  $S\hat{x}_b$  and  $S\hat{x}_m$  are efficiently separated and well reconstructed in all three cases. Furthermore, no visible deterioration arises on the images reconstructed with  $\kappa_2$ .





**Figure 11.** Evolution of the error, inside the ROI region, of the metal and tissue maps  $(Sx_{m,n})_n$  and  $(Sx_{b,n})_n$  estimated along iterations by Algorithms (11) and two choices of  $\kappa$  parameter and Algorithm (12) with  $\kappa_2$ .

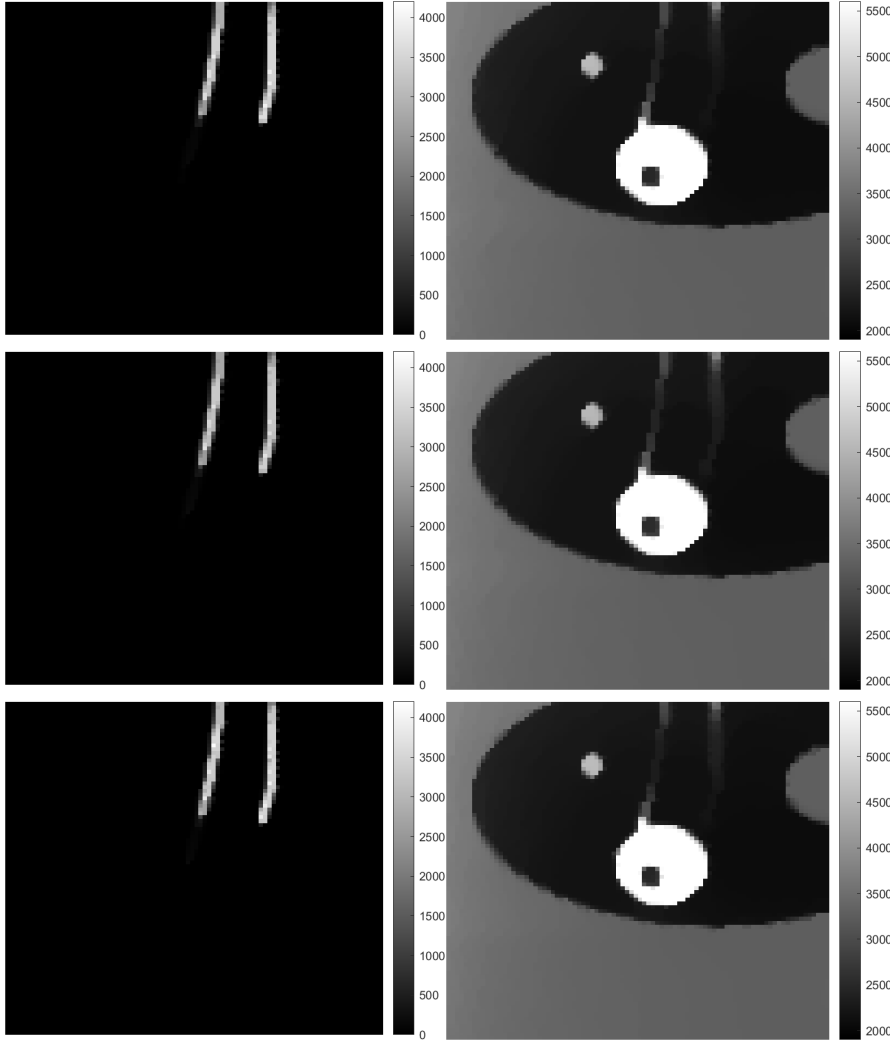


**Figure 12.** Evolution of the error, inside the ROI region, of the metal and tissue maps  $(Sx_{m,n})_n$  and  $(Sx_{b,n})_n$  estimated along iterations by Algorithm (12) with  $\kappa_1$ .

## 5. Conclusion

We have established, in a general setting, necessary conditions to ensure the convergence of the proximal gradient algorithm when the adjoint of the linear operator involved in the quadratic part of the objective function is inexact. The associated fixed point properties have been investigated. Our analysis mostly makes use of mathematical tools revolving around cocoercivity and monotone operators.

A broad class of signal recovery problems is captured by the proposed framework. The mismatched PGA can be seen as a generalization of the original PGA method. Simulations carried out in the context of X-ray tomographic imaging have shown that the proposed strategy offers an effective solution. Thus, approximate adjoints, that are, for instance, less computationally demanding, can be employed without compromising the convergence properties of the algorithm. Our theoretical results also pave the way for applications involving the refinement of the operators  $K_n$  on-the-fly or the learning of those using deep learning architectures. It would be interesting to extend our analysis to other types of data fidelity terms that may be more suitable in the presence of noise outliers such as the convex  $\ell_1$  or more robust non-convex  $\ell_p$  potentials ( $p < 1$ ) [52].



**Figure 13.** Reconstructed maps within the ROI  $S\hat{x}_m$  (left) and  $S\hat{x}_b$  (right) using  $\kappa_1$  (first row) or  $\kappa_2$  (last two rows), and either Algorithm (11) (first two rows) and Algorithm (12) (last row).

## References

- [1] Arcadu F, Stampanoni M, and Marone F 2016 On the crucial impact of the coupling projector-backprojector in iterative tomographic reconstruction *CoRR*
- [2] Bauschke H H and Combettes P L 2017 *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed
- [3] Bect J, Blanc-Féraud L, Aubert G, and Chambolle A 2004 A  $\ell_1$ -unified variational framework for image restoration *Proc. European Conf. on Computer Vision (ECCV 2004)* Prague, Czech Republic pp 1–13
- [4] Bot R I, Csetnek E R and Meier D 2019 Inducing strong convergence into the asymptotic behaviour of proximal splitting algorithms in Hilbert spaces *Optim. Methods Softw.* **34** 489–514
- [5] Briceño-Arias L M, Combettes P L, Pesquet J C, and Pustelnik N 2011 Proximal algorithms for multicomponent image processing *J. Math. Imaging Vision* **41** 3–22

- [6] Bayram I and Kamasak M E 2012 A directional total variation *Proc. European Signal Processing Conf. (EUSIPCO 2012)* Bucharest, Romania vol 19 pp 265–9
- [7] Bubba T A, Galinier M, Lassas M, Prato M, Ratti L, and Siltanen S 2020 Deep neural networks for inverse problems with pseudodifferential operators: an application to limited-angle tomography arXiv:2006.01620
- [8] Buffiere J Y, Maire E, Adrien J, Masse J P, and Boller E 2010 In situ experiments with X-ray tomography: an attractive tool for experimental mechanics *Experimental Mechanics* **50** 289–305
- [9] Candes E J, Romberg J and Tao T 2006 Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information *IEEE Trans. Inf. Theory* **52** 489–509
- [10] Chan R H, Setzer S, and Steidl G 2008 Inpainting by flexible Haar-wavelet shrinkage *SIAM J. Imaging Sci.* **1** 273–93
- [11] Chapdelaine C, Gac N, Djafari A, and Parra-Denis E 2018 New GPU implementation of separable footprint (SF) projector and backprojector: first results *Proc. 5th Int. Conf. on Image Formation in X-Ray Computed Tomography* Salt Lake City, USA pp 314–7
- [12] Chaux C, Combettes P L, Pesquet J C, and Wajs V R 2007 A variational formulation for frame-based inverse problems *Inverse Problems* **23** 1495–518
- [13] Chouzenoux E, Zolyniak F, Goullart E, and Talbot H 2013 A majorize-minimize memory gradient algorithm applied to X-ray tomography *Proc. 20th IEEE Int. Conf. on Image Processing (ICIP 2013)*, Melbourne, Australia, pp 1011–5
- [14] Combettes P L, Dung D, and Vũ B C 2011 Proximity for sums of composite functions *J. Math. Anal. Appl.* **380** 680–88
- [15] Combettes P L, Dung D, and Vũ B C 2010 Dualization of signal recovery problems *Set-Valued and Variational Analysis* **18** 373–404
- [16] Combettes P L and Pesquet J C 2007 Proximal thresholding algorithm for minimization over orthonormal bases *SIAM J. Optim.* **18** 1351–76
- [17] Combettes P L and Pesquet J C 2011 Proximal splitting methods in signal processing in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering* (H. H. Bauschke et al., eds) pp 185–212
- [18] Combettes P L and Pesquet J C 2020 Deep Neural Network structures solving variational inequalities, *Set-Valued and Variational Analysis* **28** 491–518
- [19] Combettes P L and Pesquet J C 2020 Fixed point strategies in data science arXiv:2008.02260
- [20] Combettes P L and Wajs V R 2005 Signal recovery by proximal forward-backward splitting *Multiscale Model. Simul.* **4** 1168–200
- [21] Condat L, Kitahara D, Contreras A, and Hirabayashi A 2020 Proximal splitting algorithms: Relax them all! arXiv:1912.00137
- [22] Daubechies I, Defrise M, and De Mol C 2004 An iterative thresholding algorithm for linear inverse problems with a sparsity constraint *Comm. Pure Appl. Math.* **57** 1413–57
- [23] De Man B and Basu S 2004 Distance-driven projection and backprojection in three dimensions *Phys. Med. Biol.* **49** 2463–75
- [24] Dong Y, Hansen P C, Hochstenbach M E, and Brogaard Riis N A 2019 Fixing Nonconvergence of Algebraic Iterative Reconstruction with an Unmatched Backprojector *SIAM J. Sci. Comput.* **41** A1822–39
- [25] El Gueddari L, Chouzenoux E, Vignaud A, and Ciuciu P 2019 Calibration-less parallel imaging compressed sensing reconstruction based on OSCAR regularization <https://hal.inria.fr/hal-02292372/document>
- [26] El Gueddari L, Chouzenoux E, Vignaud A, Pesquet J C, and Ciuciu P 2019 Online MR image reconstruction for compressed sensing acquisition in T2\* imaging *Wavelets and Sparsity XVIII*
- [27] Elfving T and Hansen P C 2018 Unmatched projector/backprojector pairs: perturbation and convergence analysis *SIAM J. Sci. Comput.* **40** A573–91
- [28] Figueiredo M A T and Nowak R D 2003 An EM algorithm for wavelet-based image restoration *IEEE Trans. Image Process.* **12** 906–16
- [29] Galigekere R R, Wiesent K, and Holdsworth D W 2003 Cone-beam reprojection using projection-matrices *IEEE Trans. Med. Imaging* **22**
- [30] Guay M, Czaja W, Aronova M, and Leapman R 2016 Compressed sensing electron tomography for determining biological structure *Scientific Reports* **6** <https://doi.org/10.1038/srep27614>
- [31] Herman G T 1980 *Image Reconstruction from Projections: the Fundamentals of Computerized Tomography* Academic Press

- [32] Joseph P M 1982 An improved algorithm for reprojecting rays through pixel images *IEEE Trans. Med. Imaging* **1** 192–196
- [33] Kak A and Slaney M 1988 Principles of computerized tomographic imaging *SIAM J. Sci. Comput.*
- [34] Kongskov R D and Dong Y 2018 Tomographic reconstruction methods for decomposing directional components *Inverse Probl. Imag.* **12** 1429–42
- [35] Kyrieleis A, Titarenko V, Ibison M, Connolley T, Withers P J, and Connelly T 2011 Region-of-interest tomography using filtered backprojection: assessing the practical limits *J Microsc.* **241** 69–82
- [36] Langet H, Riddell C, Reshef, Troussel Y, Tenenhaus A, Lahalle E, Fleury G, and Paragios N 2015 Compressed-sensing-based content-driven hierarchical reconstruction: Theory and application to C-arm cone-beam tomography *Med. Phys.* **42**
- [37] Liu R, Fu L, De Man B, and Yu H 2017 GPU-based branchless distance-driven projection and backprojection *IEEE Trans. Comput. Imaging*
- [38] Long Y, Fessler J, and Balter J 2010 3D forward and back-Projection for X-Ray CT using separable footprints *IEEE Trans. Med. Imaging* **29** 1839–50
- [39] Lorenz D A, Rose S, and Schöpfer F 2018 The randomized Kaczmarz method with mismatched adjoint *BIT Numer. Math.* **58** 1079–98
- [40] Moreau J J 1962 Fonctions convexes duales et points proximaux dans un espace hilbertien *C. R. Acad. Sci. Paris Sér. A Math.* **255** 2897–9
- [41] Ohnesorge B, Flohr T, Schwarz K, Heiken J P, and Bae K T 2000 Efficient correction for CT image artifacts caused by objects extending outside the scan field of view *Med. Phys.* **27** 39–46
- [42] Paleo P, Desvignes M, and Mirone A 2016 A practical local tomography reconstruction algorithm based on known subregion *J Synchrotron Radiat.* **24** 257–68
- [43] Reshef A, Riddell C, Troussel Y, Ladjal S, and Bloch I 2017 Dual-rotation C-arm cone-beam computed tomography to increase low-contrast detection *Med Phys.* **44** e164–73
- [44] Riddell C, Bendriem B, Bourguignon M H, and Kernevez J P 1995 The approximate inverse and conjugate gradient: non-symmetrical algorithms for fast attenuation correction in SPECT *Phys. Med. Biol.* **40** 269–81
- [45] Riddell C, Savi A, Gilardi M, and Fazio F 1996 Frequency weighted least squares reconstruction of truncated transmission SPECT data, *IEEE Trans. on Nuclear Science* **43** 2292–98
- [46] Leonid I. Rudin L I, Osher S, and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica D* **60** 259–68
- [47] Van Aarle W, Jan Palenstijn W, Cant J, Janssens E, Bleichrodt F, Dabrovolski A, De Beenhouwer J, Joost Batenburg K, and Sijbers J 2016 Fast and flexible X-ray tomography using the ASTRA toolbox *Opt. Express* **24** 25129–47
- [48] Xu F and Mueller K 2006 A comparative study of popular interpolation and integration methods for use in computed tomography *Proc. IEEE Int. Symp. Biomed. (ISBI 2006)*, Arlington, VA, USA, pp 1252–55
- [49] Zeng G L 2019 Counter examples for unmatched projector/backprojector in an iterative algorithm *Chinese J. Medical Imaging Technol.* 1–12
- [50] Zeng G L and Gullberg G T 2000 Unmatched projector/backprojector pairs in an iterative reconstruction algorithm *IEEE Trans. Comput. Imaging*
- [51] Zeng G L and Gullberg G T 1993 A ray-driven backprojector for backprojection filtering and filtered backprojection algorithms *Proc. IEEE Nucl. Sci. Symp. Conf. Rec. (NSSMIC 1993)*, San Francisco, CA, USA, pp 1199–201
- [52] Zeng W J, So H C, and Jiang X 2016 Outlier-robust greedy pursuit algorithms in Lp-space for sparse approximation *IEEE Trans. Signal Process.* **64** 60–75
- [53] Zhu D L and Marcotte P 1996 Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities *SIAM J. Optim.* **6** 714–26
- [54] Zou H and Hastie T 2005 Regularization and variable selection via the elastic net *J. Royal Stat. Soc. B* **67** 301–20